Design and Synthesis of Selective Kinase Inhibitors

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Karine Lafleur

aus Frankreich

Promotionskomitee

Prof. Dr. Cristina Nevado (Vorsitz und Leitung der Dissertation)

Prof. Dr. Amedeo Caflisch (Vorsitz und Leitung der Dissertation)

Prof. Dr. Kim Baldridge

Prof. Dr. Nathan Luedtke

Zürich 2012

Foreword

This PhD thesis, except for Chapter 1, is based on results published or going to be published in peer reviewed scientific journals. Chapter 2, 3, 4 and the addendum correspond to the respective manuscripts, and have been modified as little as possible. Therefore, compounds and references are numbered independently in each chapter.

Already published manuscripts include:

1. “Preparation of imidazopurinedione derivatives as antiangiogenic EphB4 kinase inhibitors useful in the treatment of angiogenesis dependent cancers and intraocular neovascular syndromes” Lafleur, K.; Huang, D.; Nevado, C.; Caflisch, A. PCT Int. Appl. WO 2011009695 A1.

2. “Complementing ultrafast shape recognition with an optical isomerism descriptor” Zhou, T.; Lafleur, K.; Caflisch, A. *J. Mol. Graph. Model*. **2010**, *29*, 443–449.

3. “Kinase Selectivity Potential for Inhibitors Targeting the ATP Binding Site: A network Analysis” Huang, D.; Zhou, T.; Lafleur, K.; Nevado, C.; Caflisch, A. *Bioinformatics* **2010**, *26*, 198-204.

4. “Structure-Based Optimization of Potent and Selective Inhibitors of the Tyrosine Kinase Erythropoietin Producing Human Hepatocellular Carcinoma Receptor B4 (EphB4)” Lafleur, K.; Huang, D. Z.; Zhou, T.; Caflisch, A.; Nevado, C. *J. Med. Chem.* **2009**, *52*, 6433-6446.

Table of Contents

Summary

Cancer is one of the leading causes of death in developed countries. During the year 2008, it affected around 7.6 million people worldwide. Chemical substances, ionizing radiation or events that lead to gene mutations are mainly responsible for the development of cancer. In particular, cancer cells are no longer able to control their normal growth and proliferation, and their accumulation induces the formation of solid tumors. Furthermore, angiogenesis involves the formation of new blood vessels supplying tumor cells with the necessary nutrients.

This PhD thesis focuses on EphB4 receptor tyrosine kinases that have recently gained attention due to their role in angiogenesis. These protein kinases play a key role in cell signaling by phosphorylating their cognate substrate using ATP as a cofactor. Thus, inhibition of kinase activity can be achieved by synthesizing small compounds that occupy the ATP binding pocket of the protein. Computer simulations are now often used to design such molecules. In particular, novel candidates for EphB4 inhibition were recently identified in our group using *in silico* high throughput screening.

The first part of this work focuses on the synthesis of these molecules and derivatives with improved biological properties. Our first goal was to optimize the interactions between the molecule and the protein in order to get a higher binding affinity. Interestingly, the inhibition activity could be improved by a thousand fold by forming additional hydrogen bonds and improving van der Waals interactions. Furthermore, the binding mode of the molecule within the ATP binding pocket was confirmed by molecular dynamics simulations. The selectivity profile of our compound was finally assessed by measuring its inhibition activity on 143 different kinases, and its high specificity proved to be comparable to currently marketed anti-cancer drugs.

However, the cellular permeability of this highly potent molecule still had to be improved. In addition, the presence of a phenol group sensitive to glucoronidation in phase II metabolism was thought to be detrimental for the pharmacological properties of the drug. Therefore, the synthesis of a second generation of EphB4 inhibitors with improved biological properties has been investigated. In addition, their interactions with the ATP binding site of EphB4 were confirmed by X-ray crystallography.

The last part of this PhD thesis describes the synthesis of new fluorescent EphB4 inhibitors. Lead optimization supported by X-ray crystallography provided a potent compound with low nanomolar inhibition activity in enzymatic assays. In addition, its biological effect on the MDA-MB-231 cell line was characterized using cellular proliferation assay and fluorescence microscopy.

Zusammenfassung

Krebs ist eine der wichtigsten Todesursachen in den entwickelten Ländern. Im Jahr 2008 hat Krebs etwa 7,6 Millionen Todesfälle weltweit verursacht. Chemische Stoffe oder Ereignisse, die zu Genmutationen oder Zelldifferenzierung führen, sind hauptsächlich für Krebs verantwortlich. Krebszellen sind Zellen, die nicht mehr in der Lage sind, ihre Vermehrung zu kontrollieren. Aus diesen unsterblichen Zellen entstehen feste Tumoren, die dank neugebildeten Blutgefäßen alle notwendigen Nährstoffe bekommen. Dieser Prozess wird Angiogenese genannt.

Diese Dissertation fokusiert auf EphB4 Rezeptor-Tyrosinkinasen, die in den letzten Jahren wegen ihrer Rolle bei der Angiogenese Aufmerksamkeit gewonnen haben. Protein Kinasen richten Signale aus, indem sie eine Phosphatgruppe von einem ATP Molekül auf ein Substrat transferieren. Um die Aktivität von Kinasen zu inhibieren, können kleine Verbindungen hergestellt werden, die die ATP Bindungstasche von dem Enzym füllen. Um solche Inhibitoren zu entwickeln, werden Computersimulationen oft gebraucht. Im Jahr 2008 wurde in unserer Forschungsgruppe ein neuer EphB4 Inhibitor durch Computerrechnungen entdeckt.

Der erste Teil dieser Arbeit konzentriert sich auf die Synthese von ähnlichen Derivaten, die EphB4 besser inhibieren sollten. Das Ziel war es, die Interaktionen zwischen dem Molekül und dem Protein zu optimieren, um eine höhere Aktivität zu bekommen. Interessanterweise wurde der Inhibitor tausend Mal verbessert, indem zusätzliche Wasserstoffbrückenbindungen gebildet wurden. Ausserdem wurden die Interaktionen zwischen dem Molekül und dem Protein durch Computersimulationen bestätigt. Um Nebenwirkungen einzuschätzen, wurde der Inhibitor schliesslich auf 143 anderen Kinasen getestet, und sein selektives Profil sah ähnlich oder besser aus als andere verkaufte Medikamente.

Diese sehr aktive Verdindung musste aber noch verbessert werden, da ihre Zellpermeabilität noch nicht optimal war. Ihre Phenolgruppe könnte sogar andere Probleme verursachen, indem das Molekül schneller metabolisiert werden könnte. Im zweiten Kapitel wird deswegen die Synthese von einer zweiten Generation von EphB4 Inhibitoren beschrieben, die bessere pharmakokinetische Eigenschaften haben. Ihre biologischen Aktivitäten wurden auf verschiedenen Zelllinien getestet, und ihre Interaktionen mit dem Protein wurden durch Röntgenstrahlungen bestätigt.

Im dritten Teil wird die Synthese von neuen EphB4 Inhibitoren beschrieben, die sehr interessante Fluoreszenzeigenschaften haben. Ihre biologische Aktivitäten wurden auf MDA-MB-231 Zellen getestet, und ihre optimale Absorption- und Emissions-Wellenlängen wurden für Zellbildgebung genutzt.

Chapter 1: Introduction

1 Introduction

The introduction of this dissertation will cover several topics related to anti-cancer therapies. In particular, the inhibition of the tyrosine kinase EphB4 (erythropoietin producing human hepatocellular carcinoma receptor B) will be presented as an attractive strategy for the development of new anti-angiogenic therapies. Next, the process of drug discovery following the identification of target receptors will be described, and the role of computer aided drug design will be particularly emphasized. Finally, the characterization of a drug mode of action using fluorescent molecules and cellular imaging techniques will be briefly discussed.

1.1 Cancer: definition and target identification

1.1.1 Definition

Cancer is one of the most important causes of death in developed countries. In year 2008, it affected about 12.7 million people, and caused about 7.6 million deaths worldwide.[1] Cancer cells are defined as cells that have lost the regulatory mechanism controlling their growth and proliferation. The disease is considered as benign when it stays localized, and becomes malignant as soon as abnormal cells invade other tissues. Chemical substances or events that induce gene mutations or interfere with cell differentiation are mainly responsible for the development of cancer.

Defects induced by genetic mutations include modifications of signaling pathways that regulate cell growth and division.[2] In addition, apoptosis, a process that controls cell death, can be overcome upon down-regulation of apoptosis-inducing proteins, such as p53. In order to have sufficient blood supply and obtain the necessary nutrients, tumor cells release growth factors that activate the formation of new blood vessels. This process, called angiogenesis, usually occurs in the body to repair injured tissues, but is uncontrolled in the case of tumor growth. Finally, down-regulation of cell adhesion molecules allows malignant cells to leave the primary tumor and invade other tissues. This stage is called metastasis.

1.1.2 Anti-cancer agents

Several examples of chemotherapeutic drugs

Anti-cancer agents targeting DNA:

Examples include planar molecules, so called intercalators, which fit between the base pairs of DNA, or alkylating agents able to form covalent bonds with the nucleophilic centers of DNA. Once bound, these molecules inhibit the replication and transcription of DNA. Although intercalators and alkylating agents can lead to side effects as they also target DNA in healthy cells, molecules like cisplatine[3] (Figure

1) have been shown to be useful in the treatment of cancer. Indeed, it is thought that the inhibition of replication and transcription of DNA mostly affect tumor cells that divide faster than normal cells. In addition, the biosynthesis of DNA and its nucleic acids can be inhibited using antimetabolites which bind to the enzymes involved in the process.

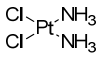


Figure 1. Alkylating agent cisplatin.

Anti-cancer agents targeting tubuline:

Tubuline is a structural protein that polymerizes to form microtubules essential to maintain the shape of the cell. Microtubules also play an important role during cell division, as they depolymerize back to tubulin, which acts as a building block to form mitotic spindles (Figure 2, adapted from Ref. 2).

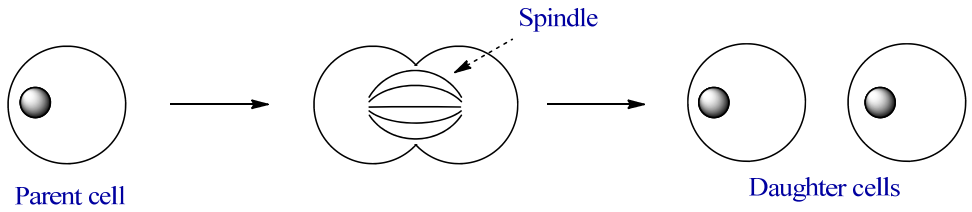


Figure 2. Cell division.

Drugs inhibiting this process upon interaction with tubulin represent a second class of anti-cancer agents. One of the best known examples is paclitaxel (taxol, Figure 3).[4]

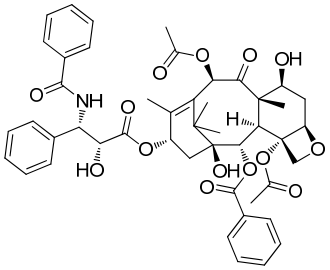


Figure 3. Tubulin inhibitor paclitaxel.

Protein kinase inhibitors

The selectivity of anti-cancer agents remains one essential challenge in drug discovery. Indeed, targeting proteins that are only expressed or over-expressed in cancer cells would greatly reduce the risk of side effects. For example, certain protein kinases have been shown to be present in larger quantities in many cancer cells. As kinases regulate key processes such as cell growth and cell division, their inhibition has been recognized as an attractive strategy for the treatment of cancer.

Receptor tyrosine kinases (RTKs) are transmembrane proteins composed of a C-terminal cytoplasmic protein tyrosine kinase (PTK) domain, a N-terminal domain necessary for ligand binding and a transmembrane segment.[5] Binding of RTKs to their cognate ligands induces receptor dimerization and subsequent autophosphorylation (Figure 4). Activated kinases regulate the activity of cytoplasmic proteins upon tyrosine phosphorylation. Interaction with protein phosphatases induces hydrolysis of the phosphorylated residues and switches off the activity of the protein tyrosine kinase.

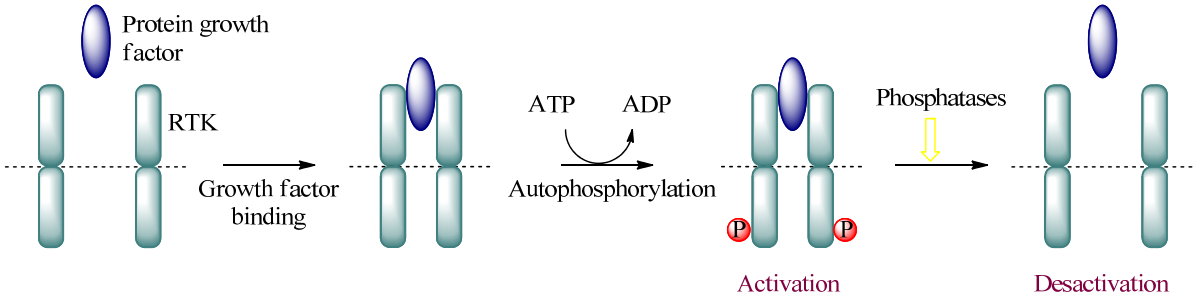


Figure 4. Dimerization and activation/inactivation of receptor tyrosine kinases.

Activation of the protein kinase activity induces a signaling cascade that ultimately controls transcription of specific genes in DNA (Figure 5, adapted from Ref. 5).[5] As an example, the Ras-activated MAP kinase cascade involves binding to the Grb2-SOS complex, which in turn, induces the G-protein Ras to exchange its bound GDP for GTP. Activation of Raf kinases upon interaction with Ras·GTP promotes activation of MEK kinases able to phosphorylate MAP kinases. MAP kinases activate other proteins and, altogether, enter the nucleus of the cell to phosphorylate transcription factors such as Fos, Myc, and Jun proteins. This final stage induces alterations in gene expression.

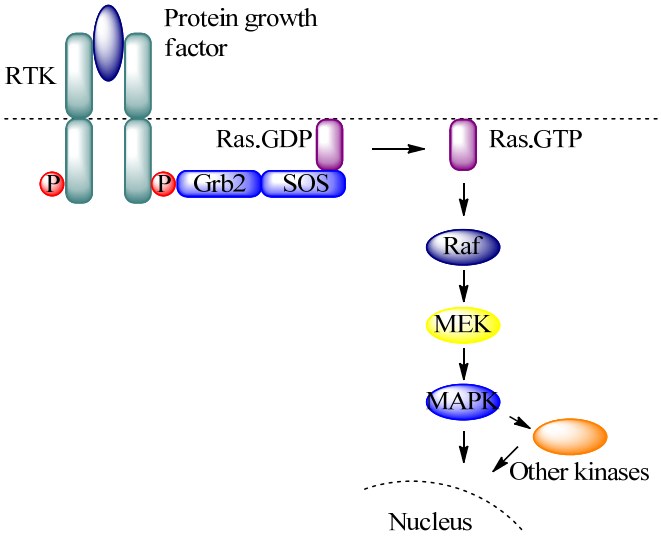


Figure 5. Ras-activated MAP kinase cascade.

Adenosine triphosphate (ATP) acts as a phosphorylating agent and binds to the active site located between the N-terminal and the C-terminal lobes of the kinase domain. The ATP binding pocket can be divided into five different subpockets (phosphate binding pocket, sugar region, entrance pocket, adenine binding region and hydrophobic region) as shown in Figure 6. The hydrophobic pocket, which is located at the opposite site of the ribose pocket, is not occupied by ATP and proves to be very important in the design of selective inhibitors.[6] The accessibility to this pocket depends on the size of the so-called gatekeeper residue.

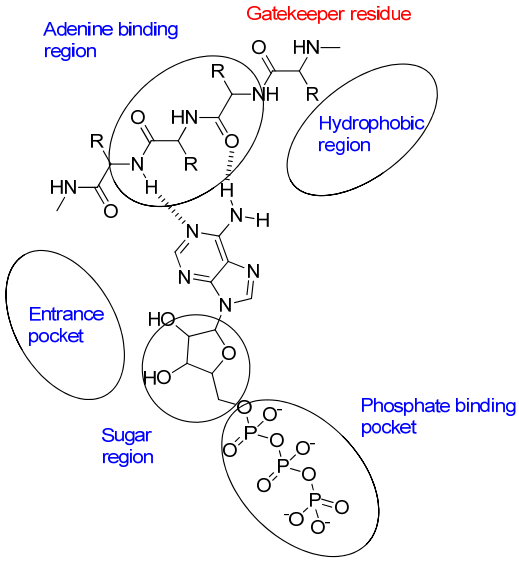


Figure 6. Binding map of ATP bound to the catalytic site of a kinase.

The design of ATP competitive inhibitors mainly started in 1986 with the discovery of staurosporine as a nanomolar inhibitor of the protein kinase C.[7] Later, many small molecules reached the market, including dasatinib (Figure 7, left),[8] a Bcr-Abl inhibitor used in the treatment of chronic myeloid leukemia (CML). However, as every protein kinase uses ATP as a phosphorylating agent, the binding site is highly conserved from one protein to another, and the design of selective kinase inhibitors still remains a challenge. Recently, the development of type II inhibitors has emerged as an alternative strategy to improve selectivity.[9] Type II inhibitors –as opposed to type I inhibitors– bind to the inactive, non-phosphorylated conformation of the protein. Their selectivity profile is enhanced as they occupy a hydrophobic pocket that is only exposed when the kinase adopts an inactive conformation (Figure 7, right). One of the first examples of type II inhibitors is imatinib,[10] which was approved for clinical use in 2001.

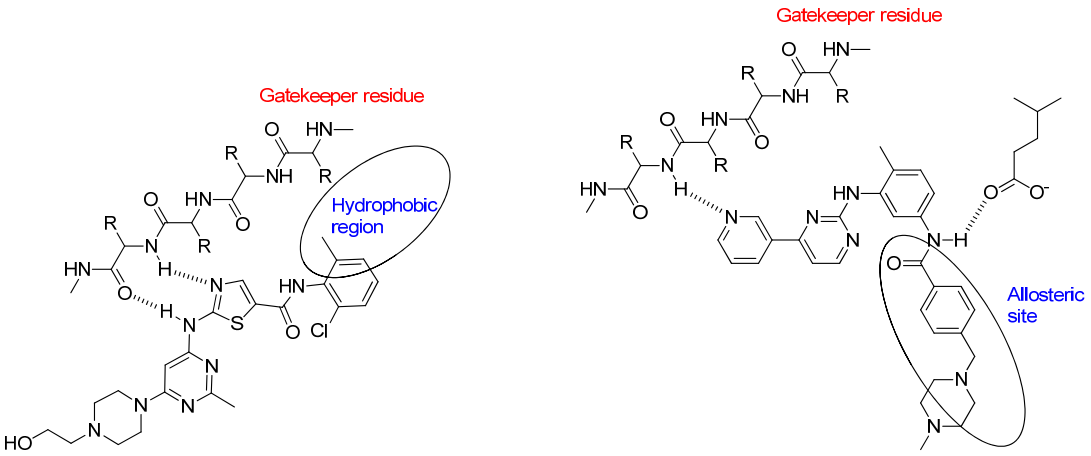


Figure 7. Kinase bound to dasatinib in its active form (left) and kinase bound to imatinib in its inactive form (right).

1.1.3 EphB4, an attractive target for anti-cancer therapy

Eph receptor tyrosine kinases are the largest family of RTKs. They include 14 receptors and 8 ligands that have been divided into two subclasses, A and B.[11] In general, EphA receptors bind to ephrinA ligands, while EphB receptors bind to ephrinB ligands. Both EphA and EphB receptors are transmembrane proteins. Their extracellular region is composed of an N-terminal domain necessary for ligand binding, whereas their intracellular domain includes a C-terminal domain and a tyrosine kinase domain (Figure 8).

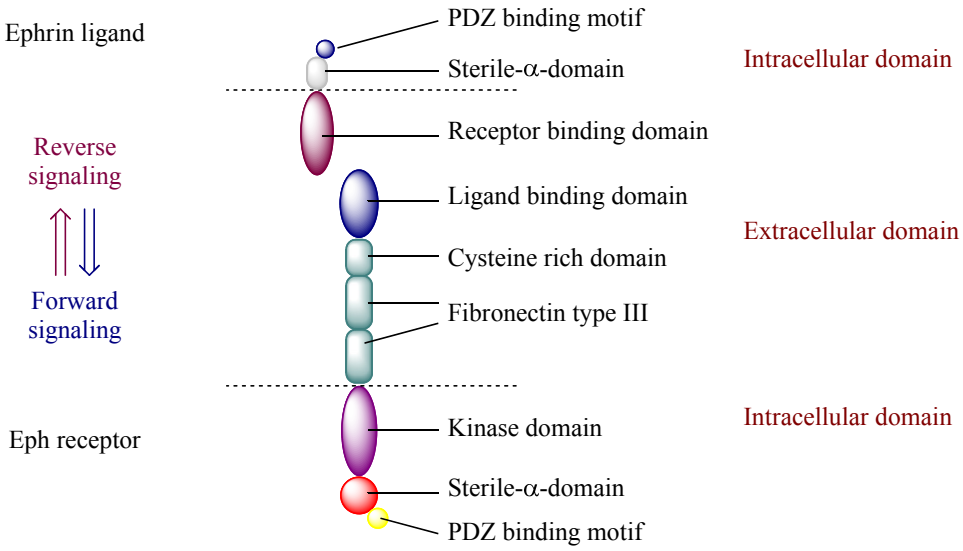


Figure 8. Structure of ephrin ligands and Eph receptors

Eph-ephrin interaction is a complex event that induces a bidirectional signaling: a receptor forward signaling propagating into the receptor-expressing cell and a ligand reverse signaling propagating into

the ligand-expressing cell. While other receptor tyrosine kinases regulate cell proliferation and differentiation, Eph-ephrin signaling controls cell migration, repulsion, adhesion or attachment to the extracellular matrix. Eph-ephrin interaction can promote a repulsive response through a protease-mediated cleavage of the Eph or ephrin extracellular domain, or through internalization of the Eph-ephrin complex into the receptor or ligand-expressing cell. On the other hand, prolonged Eph-ephrin signaling can induce cellular adhesion.[12]

Multiple aspects of the role of Eph-ephrin signaling in cancer have been reported.[12] First, Eph forward signaling has been shown to inhibit proliferation and migration of many cancer cells. Cellular contact between cancer cells is favored, and invasion of normal tissues is avoided. These observations are consistent with low ephrin expression or Eph mutations in cancer cells that reduce Eph-ephrin contact.[13] However, in some cases, Eph expression can promote cancer progression. In fact, down-regulation of EphB4 or EphA2 by antisense oligonucleotides has been shown to decrease cancer progression in several cancer models. Interaction with other signaling proteins rather than ephrin stimulation or kinase activity might be responsible for the oncogenic ability of Eph receptors.[12]

In another context, EphB4-ephrinB2 signaling contributes to blood vessel remodeling during vascular development. Treatment with soluble extracellular domains of EphB4 has been shown to inhibit tumor growth and angiogenesis in xenograft studies.[14,15] In addition, a recent study indicated that suppression of Eph forward signaling using the kinase inhibitor NVP-BHG712 inhibits Vascular Endothelial Growth Factor (VEGFR) driven angiogenesis.[16] However, EphB4 seems to be rather recalcitrant to inhibition since, despite its potential therapeutic importance, only a few series of non-peptidic small molecule inhibitors have been reported in the literature to date (Figure 9).[8,17,18]

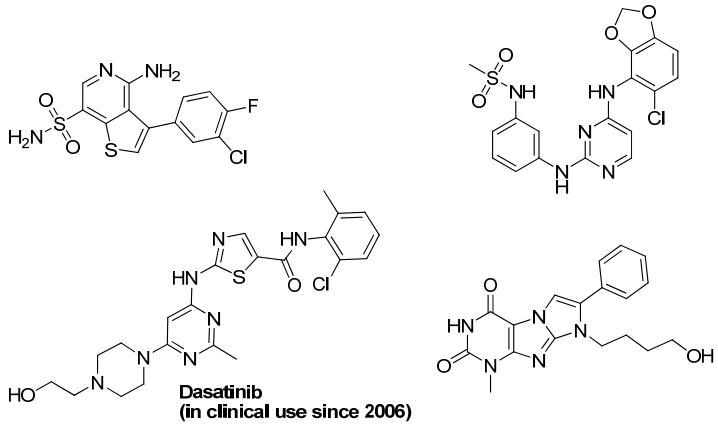


Figure 9. Known EphB4 inhibitors

1.2 Computer aided drug discovery

1.2.1 Drug discovery process

The first stage of a drug discovery process represents the identification of a target receptor involved in a specific disease area (Figure 10). Next, libraries of synthetic compounds or natural products are tested ("screened") on the target (DNA, protein, etc…), either experimentally or computationally. The isolation of natural products from plants or microorganisms has provided many biologically active compounds, such as taxol[4] or vancomycin.[19] In addition, other compounds such as synthetic intermediates, modified drugs or natural ligands are regularly included in screening campaigns.

Virtual screening and *de novo* design techniques (section 1.2.2.) are commonly used to select molecules for *in vitro* screenings upon analysis of the receptor shape and interaction sites. If the structure of the target protein has not been solved by X-ray crystallography, alternative strategies may be considered, such as the design of homology models from similar proteins or the use of NMR spectroscopy.

In vitro assays are relatively easy to establish, as they involve isolated enzymes or cells and can be automated, enabling the screening of thousands of compounds in a very short time (High Throughput Screening). In addition, sufficient amount of enzymes can be easily produced by fast-growing cells such as bacteria. Positive hits on enzymatic assays are further tested on cellular based assays, where compound binding and pharmacokinetic properties can be assessed upon cell proliferation or colorimetric measurements.

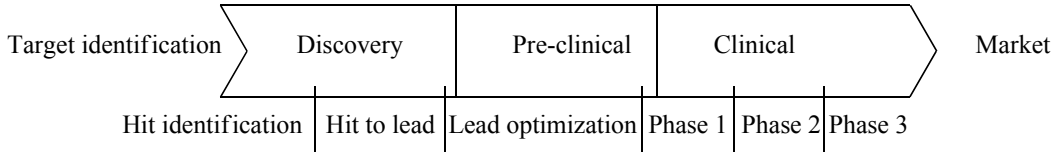


Figure 10. The drug discovery process

Molecules binding to the target and displaying the desired pharmacological effects are designated as "Hits". These compounds may be re-synthesized and modified to improve drug-like properties such as water solubility, selectivity, and cell membrane permeability. Further biological evaluations involve *in vivo* assays requiring specific animal models. For example, the anti-inflammatory effect of a drug is observed upon induction of inflammation in mice. Compounds showing the desired therapeutic effect on animal models are selected for preclinical and clinical studies where safety and efficacy will be later on assessed on humans.

1.2.2 Computational tools in drug design

Computational tools can be divided into two main families: quantum mechanics and molecular mechanics.[20] In molecular mechanics, electrons are not taken into account explicitly, and nuclei and bonds are considered as balls and springs, respectively. Equations derived from classical physics are integrated to evaluate bonding (stretching, bending, torsion) and non-bonding (van der Waals, electrostatics) interaction energies (Equation 1). Calculations with molecular mechanics are usually fast and suitable for relatively big molecules, such as proteins.

$$E^{total} = \sum_{i}^{bonds} E_i^{stretch} + \sum_{i}^{angles} E_i^{bend} + \sum_{i}^{dihedrals} E_i^{torsion} + \sum_{ij}^{atompairs} E_{ij}^{vdW} + \sum_{ij}^{atompairs} E_{ij}^{elec}$$

$$E^{total} = \sum_{i} \frac{1}{2} k_i^r (r_i - r_i^0)^2 + \sum_{i} \frac{1}{2} k_i^\theta (\theta_i - \theta_i^0)^2 + \sum_{i} \frac{1}{2} k_i^T (1 + \cos(n_i \tau_i - \delta_i)) + \sum_{ij} \left(\frac{-A_{ij}}{r_{ij}^6} + \frac{B_{ij}}{r_{ij}^{12}}\right) + \sum_{ij} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}$$

Equation 1. Energy in molecular mechanics. The total energy is the sum of stretching (black) and bending (red) terms, torsion terms (blue) as well as van der Waals (green) and electrostatic (purple) energies. k_i^r, k_i^θ, and k_i^T are the bond, angle and dihedral angle force constants. r_i, θ_i and τ_i are the bond lengths, angles and dihedral angles.

Quantum mechanical methods solve the Schrödinger equation (Equation 2), enabling the calculation of structure and electronic properties, such as molecular orbitals and energies, dipole moments, and partial charges. Integration methods can be divided into two categories. *Ab initio* methods are very rigorous as they involve the exact evaluation of integrals, but are time demanding and restricted to small molecules. On the other hand, semi-empirical approaches, such as AM1 or PM3, neglect particular integrals and include experimentally derived parameters. Thus, these methods are faster and can be applied to larger systems.

$$\hat{H}\varphi = E\varphi$$

$$\hat{H} = -\sum_{\alpha} \frac{\hbar^2}{2} \frac{1}{M_\alpha} \nabla_\alpha^2 - \sum_{i} \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_{\alpha} \sum_{i} \frac{Z_\alpha e^2}{4\pi\varepsilon_0 r_{i\alpha}} + \sum_{i} \sum_{j} \frac{e^2}{4\pi\varepsilon_0 r_{ij}} + \sum_{\alpha} \sum_{\beta} \frac{Z_\alpha Z_\beta e^2}{4\pi\varepsilon_0 r_{\alpha\beta}}$$

Equation 2. Schrödinger equation. The hamiltonian is the sum of nuclear and electronic kinetic energies (red) and electrostatic potentials (blue).

Energy minimization and conformational search

Geometry optimizations usually represent the first step of a computational study. Here, the potential energy of the molecule is represented by a hypersurface where each point corresponds to the energy of

a particular structure (Figure 11). Geometry optimization algorithms expand energy functions as Taylor series (Equation 3), and can be classified into order 0 (grid search), order 1 (steepest descent and conjugate gradient) and order 2 algorithms (Newton Raphson) depending on the highest derivative they use.

$$E(x) = E(x_0) + (x - x_0)E'(x_0) + {}^1/_2\,(x - x_0)^2 E''(x_0) + \cdots$$

Equation 3. Expansion of energy functions as Taylor series.

However, energy minimizations usually generate a structure close to the geometry that has been entered. Thus, in order to explore the energy surface and find the global minimum, the system has to overcome a certain energy barrier.[20] In molecular dynamics (see below), the system follows Newtons equations of motions, and a variety of different conformations can be generated. Other approaches, such as Monte Carlo and Metropolis methods, randomly modify bond lengths and angles and may be run at different temperatures. Genetic algorithms encode the conformational information of a given molecule into a chromosome, and a new set of conformations is created upon mutations of the genetic information. Finally, a systematic rotation of rotatable bonds allows a complete exploration of the energy surface, but the number of generated structure dramatically increases with the size of the molecule (for example, almost 36 million structures should be generated for a systematic rotation of seven rotatable bonds using an increment of 30°).[20]

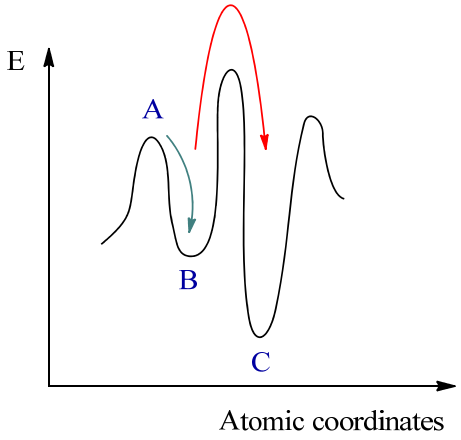


Figure 11. Conformational search. Starting from structure A, a simple energy minimization will generate conformation B, which is a local minimum. However, a conformational search will generate global minimum C.

Docking

Protein structures obtained from the Protein Data Bank must be slightly modified to be suitable for virtual screening campaigns. For example, the protonation state of histidine residues must be carefully

assigned. In addition, residues bearing a carboxylic group such as glutamic and aspartic acids must be negatively charged. Other amino acids having a primary amine or a guanidinium moiety such as lysines or arginines, respectively, are positively charged. Hydrogen bonds must be added and care should be taken upon addition of polar hydrogens that may adopt different conformations. Finally, atom types and partial charges must be assigned, and the protein can be shortly minimized.

Virtual screening campaigns have been recognized as a powerful strategy to select molecules for biological screenings. Usually, a library of compounds is docked in every positions of the binding site, such that optimal orientations can be determined. A score is assigned to each pose, and the best-scored molecules are selected. As an example, the docking program Autodock[21] covers the binding site of the protein with a grid. A probe atom is placed at each intersection point, and affinity potentials with the receptor are evaluated and stored in look-up tables. Depending on the position of the molecule within the binding site, interaction energies can be quickly evaluated upon addition of the previously listed values.

An alternative strategy is based on a preliminary fragmentation of molecular structures. In the group of Prof. Caflisch, a program named DAIM[22] has been developed which first identifies and breaks non terminal rotatable bonds within the molecule (Figure 12). DAIM generates a fingerprint that describes the chemical features of each fragment, such as atom types, bond types, and number of rings. Molecular fragments with the highest chemical richness (sum of all entries in the DAIM fingerprint) are docked in the binding site using a second program called SEED, and final poses are clustered together.[23] Finally, FFLD[24] uses the best poses obtained by SEED to optimize the orientation of the whole molecule within the pocket. Energy minimization of the final poses is performed by the program CHARMM.[25]

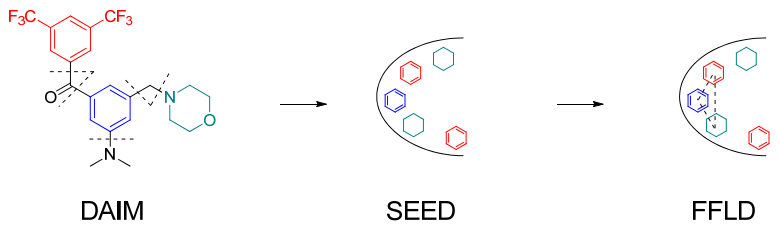


Figure 12. The DAIM/SEED/FFLD process. The fragments generated by DAIM are docked within the binding site of the protein using SEED. FFLD connects the fragments together to find the best pose.

De novo design

While virtual screening campaigns rely on already known compounds, *de novo* design programs aim at creating novel molecular structures.[26] If the structure of the target protein is already known, residues within the binding site that may be involved in hydrogen bonding or van der Waals interactions are

identified, and molecules able to interact with them can be designed. However, if the structure of the protein has not been determined, interaction sites and important functional groups may be predicted from a set of already known inhibitors.

De novo design strategies can be broadly divided into linking and growing approaches. In a linking approach, hydrogen bond donors and acceptors are identified within the binding site as well as residues involved in van der Waals interactions.[27] Molecular fragments are positioned within the pocket so that their functional groups overlap the available interaction sites. Finally, fragments predicted to be involved in favorable van der Waals and hydrogen bonding interactions with the receptor are identified and linked together using a stored library of connecting bridges (Figure 13).

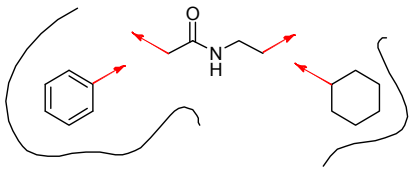


Figure 13. Linking approach. A library of fragments is docked in the binding pocket and promising poses are linked together.

The second approach, the growing program, first covers the binding pocket with a grid.[28] A probe atom is placed at each intersection point and interaction energies with the binding site are calculated. The growing process starts from atoms or fragments involved in hydrogen bonds with the binding pocket. A molecular skeleton is grown one atom at a time upon sequential addition of randomly chosen atom types, bond types and torsion angles. At each stage of the process, the steric energy of the growing molecule is evaluated. Molecular structures involved in favorable interaction energies with the binding site are further grown until they reach a size defined by the operator.

One of the disadvantages of *de novo* design is the potential identification of molecules that are difficult or impossible to synthesize. Thus, a few *de novo* design programs also include a retrosynthetic analysis. In particular, connection rules have been implemented that derive from organic synthesis reactions or statistical knowledge based on the occurrence of chemical bonds in existing molecules.[26,29,30]

Scoring functions

Docking and *de novo* design campaigns usually generate a considerable amount of possible hits. Therefore, scoring functions have been developed in order to discard unrealistic poses. Importantly, the evaluation of scoring functions must be fast and efficient, and should be able to reproduce experimental

results. In particular, scoring functions should rank a set of inhibitors in a reliable way, such that active compounds may be selected.

Scoring functions have been divided into three different categories.[31] First, force field based scoring functions are based on a combination of van der Waals and electrostatic contributions. GANDI[32] evaluates van der Waals energies using a 6-12 Lennard-Jones potential, and electrostatic energies with a distance dependant dielectric model. Both inter and intramolecular interactions are taken into account (Equation 4). Force field based scoring functions tend to overweight favorable ionic interactions, whose values can be three to four times higher than total van der Waals interactions.

$$E_{ff} = E_{vdW}^{inter} + E_{elec}^{inter} + E_{vdW}^{intra} + E_{elec}^{intra}$$

Equation 4. Force field scoring function in GANDI.

Empirical scoring functions correlate experimental binding energies (ΔG_{bind}) to empirical functions that are based on a weighted sum of entropic, hydrophobic and hydrogen bonding contributions. Individual weights are determined upon multiple linear regression. LIECE[33] is a combination of force field and empirical scoring function, as it correlates the sum of van der Waals and electrostatic interactions to the experimental binding energies (Equation 5). Each coefficient can be evaluated using a set of known inhibitors.

$$\Delta G_{bind} = \alpha \Delta E_{vdW} + \beta \Delta G_{elec}$$

Equation 5. LIECE scoring function.

Knowledge based scoring functions are based on statistical analyses derived from experimentally known structures. Interatomic distances are considered as favorable if they have been frequently observed. The advantage of knowledge based function is the implicit consideration of entropic and desolvation effects.

Molecular dynamics simulations

Molecular dynamics simulations monitor atomic displacements over a certain period of time and are very useful in conformational search and protein function studies. Solvent effects may be implicitly taken into account upon introduction of an effective dielectric constant within the scoring function. To be more accurate, molecules of solvent can be explicitly added around the protein. Usually, the macromolecule is then immersed in a water box and the molecules of solvent overlapping the protein are deleted (Figure 14).

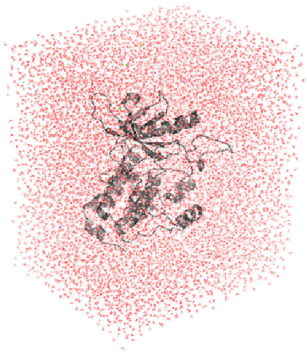


Figure 14. Molecular dynamics simulation in explicit water.

In order to simulate an infinite system, the cubic box is periodically repeated in every direction. Water molecules are minimized before the simulation starts, while keeping the protein structure rigid (Figure 15).

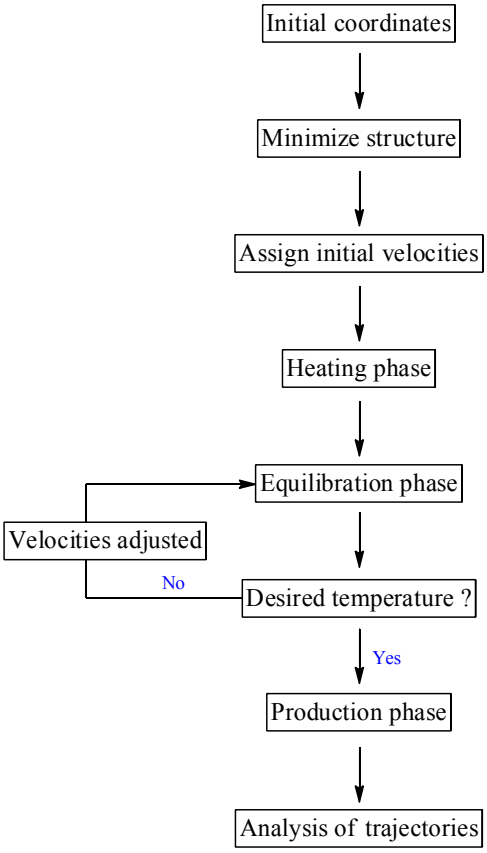


Figure 15. MD simulation process.

Initial velocities are assigned to each atom according to a Maxwell-Boltzmann distribution and the forces acting on them are calculated. The corresponding accelerations are then evaluated upon

integration of Newton's laws of motion so that positions and velocities at the next time step can be predicted. During the heating phase, the temperature T of the system increases gradually and the atomic velocities are reassigned according to the equipartition theorem:

$$\frac{3}{2}NkT = \sum_{i=1}^{N} \frac{m_i v_i^2}{2}$$

where N is the number of atoms, k is the Boltzmann constant, m_i is the mass of atom i and v_i the velocity of atom i.

Once the desired temperature has been reached, the simulation continues until properties such as temperature, pressure and energy become stable (equilibration phase). If one of these properties varies significantly, the velocities are rescaled such that it returns to its desired value. Finally, the simulation is run for the time desired (production phase).

In addition to atomic trajectories, important properties such as average energies (mean energy: $\langle E \rangle = \frac{1}{N}\sum_{i=1}^{N} E_i$, where N is the number of steps) or RMSDs can be calculated at the end of a molecular dynamics simulation. Root mean squared distances (RMSDs) are particularly useful to compare two objects upon superimposition and evaluate the stability of a protein structure during the simulation. RMSDs are calculated as following:

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - x_{i,ref})^2 + (y_i - y_{i,ref})^2 + (z_i - z_{i,ref})^2}$$

where N is the number of atoms, $x_{i,ref}$, $y_{i,ref}$, and $z_{i,ref}$ represent the coordinates of the structure of reference and x_i, y_i and z_i represent the coordinates of the structure to evaluate.

Thus, MD simulation is a powerful tool to predict time-dependent behaviors of molecular systems.

1.3 Development of fluorescent drugs

1.3.1 Basic principles of fluorescence

The absorption of photons involves the promotion of electrons to unoccupied orbitals, and is followed by radiative and non-radiative de-excitation processes. In particular, photoluminescence is related to the emission of photons upon relaxation of electronically excited molecules, and include fluorescence, phosphorescence and delayed fluorescence. Before discussing the multiple applications of fluorescence, basic principles of electronic transitions will be underlined here.

First, the combination of s/s, s/p_z or p_z/p_z orbitals leads to the formation of σ bonding and σ* non-bonding orbitals. In addition, π bonding and π* non bonding orbitals are formed upon orbital overlap of

p_x/p_x or p_y/p_y orbitals. Non bonding electrons on heteroatoms are located in n orbitals (Figure 16). The multiplicity M of the system is given by M = 2S+1, where S is the total spin quantum number ($S = \Sigma s_i$ with $s_i = \pm 1/2$). Singlet states (M = 1) are denoted by S_i, where i is the level of the excited state. Triplet states (M = 3) are denoted by T_i and have lower energies than singlet states according to Hund's rules.

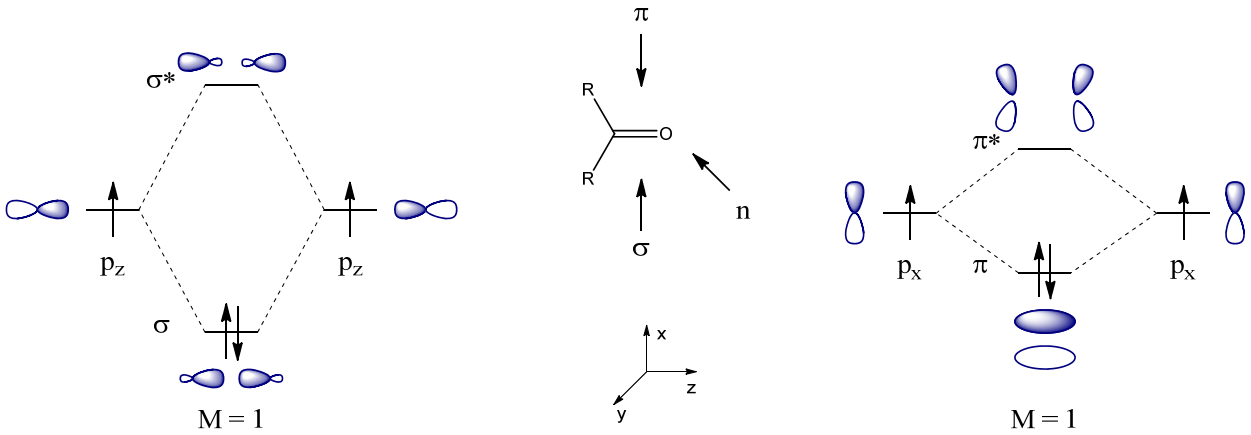


Figure 16. Formation of σ and π bonds.

In order to predict electronic transitions, a dipole moment representing the displacement of the charges during the transition is introduced:

$$\mu_{12} = \langle \varphi_1 | \mu | \varphi_2 \rangle$$

where φ_1 and φ_2 represent the wavefunctions of states 1 and 2 and μ represents the dipole moment operator. The probability of transition P_{12} is proportional to $|E.\mu_{12}|^2$, where E corresponds to the electric vector, and electronic transitions will be allowed if $P_{12} \neq 0$. For example, transitions between states of different multiplicities are forbidden.[34]

1.3.2 Perrin-Jablonski diagram

De-excitation processes have been underlined on the Perrin-Jablonski diagram (Figure 17).[34] Internal conversion stands for non-radiative transitions between states of identical spin multiplicity. On the other hand, radiative transitions between S_1 and S_0 correspond to fluorescence emissions usually occurring after non-radiative relaxations to the lowest vibrational level. Therefore, the emission spectrum is shifted to higher wavelength and the "Stokes shift" refers to the difference between the maximum wavelength of absorption and the maximum wavelength of emission (Figure 18).

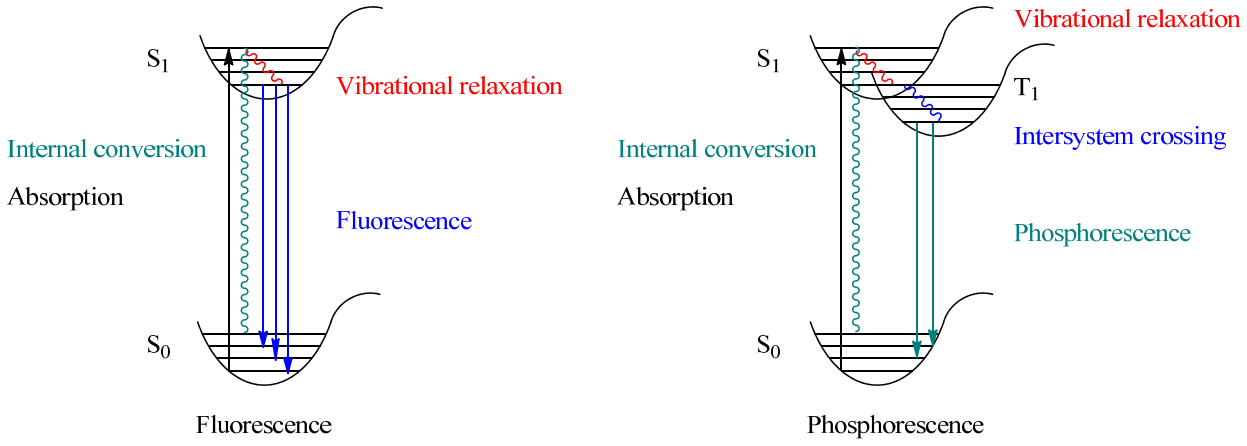


Figure 17. Perrin-Jablonski diagram.

Intersystem crossing is related to non-radiative transitions between states of different multiplicities. These forbidden transitions may be allowed in the presence of heavy atoms favoring a large spin orbit coupling. In addition, radiative phosphorescence emission can be observed at low temperature, where molecular collisions allowing intersystem crossing are avoided. Finally, delayed fluorescence may occur after reverse intersystem coupling between T_1 and S_1.

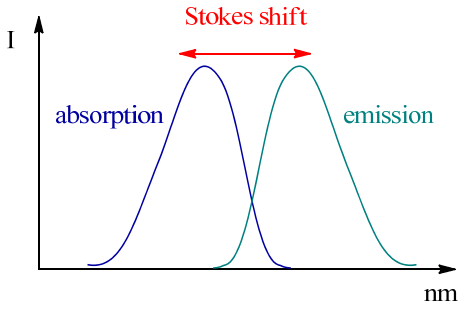


Figure 18. Stokes shift

1.3.3 Excited life time and quantum yields

Fluorescent molecules are characterized by their excited-state life times, which define the observation time window of dynamic processes.[34] The relaxation rate from the excited state can be written as:

$$-\frac{d[A^*]}{dt} = (k_r + k_{nr})[A^*]$$

where [A*] is the concentration of molecules in the excited state S_1. k_r and k_{nr} represent the radiative and non-radiative transition rate constants. The time evolution of [A*] is obtained upon integration of the relaxation rate:

$$[A^*] = [A^*]_0 \exp\,[-t(k_r + k_{nr})] = [A^*]_0 \exp\,[\frac{-t}{\tau}]$$

where $[A^*]_0$ is the initial concentration of excited molecules. The lifetime of the exited state is defined as:

$$\tau = \frac{1}{k_r + k_{nr}}$$

The fluorescence quantum yield φ is the ratio of emitted photons to absorbed photons and is given by:

$$\varphi = \frac{k_r}{k_r + k_{nr}} = k_r \tau$$

1.3.4 Molecular structure and fluorescence

Although molecular fluorescence properties are relatively difficult to predict, a few experimental observations can be mentioned.[34] First, it is important to note that exited-state lifetimes associated with n => π* transitions are higher than those associated with π => π* transitions, thus favoring non-radiative relaxations (Figure 19). For example, heterocyclic molecules or nitro-substituted compounds are usually poorly fluorescent due to low lying n => π* transitions. However, hydrogen donating solvents ("protic solvents") interacting with the nitrogen atom may invert the transition order and improve the fluorescence properties of the molecule. In contrast, lone pairs of electron donating groups, such as alcohols or amines, are involved in π bonding with the aromatic system, and induce a red shift in the absorption and emission spectra. Finally, the presence of heavy atoms often favors a large spin orbit coupling, thus enabling intersystem crossing and fluorescence quenching.

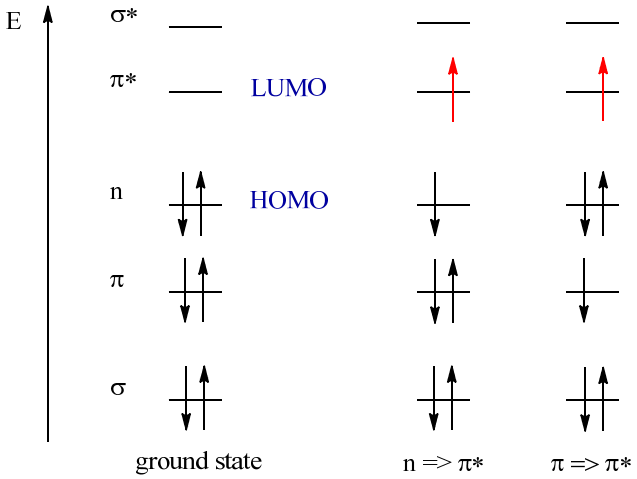


Figure 19. Illustration of π , n and π* energy levels. Solvation of the lone pair in polar environment decreases the energy level of the n orbital, and can induce an inversion of the lowest-lying transitions (n => π* vs. π => π*).

1.3.5 Development of fluorescent dies as biological imaging tools

The combination of a green fluorescent protein (GFP) with a protein of interest has emerged as a powerful strategy to understand protein function and protein distribution using fluorescence microscopy.[35,36]

The development of fluorescent probes represents an alternative strategy for the characterization of biological processes, as fluorescence emission is extremely sensitive to the microenvironment of the molecule (solvent polarity, viscosity and acidity). Common synthetic approaches usually involve covalent binding of fluorophores to a given scaffold. For example, coupling of purvalanol B to a dansyl moiety yielded potent cyclin-dependent kinase (CDK) inhibitors, where cellular distribution could be characterized using fluorescence microscopy (Figure 20).[37]

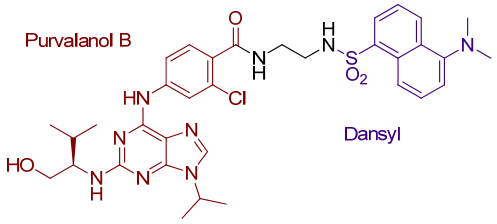


Figure 20. Development of fluorescent CDK inhibitors.[37]

In many cases, fluorophore-conjugated molecules may present lower potency and higher toxicity than the original scaffold. In addition, cleavage of covalent bonds within the intracellular environment potentially leads to an incorrect interpretation of cellular distribution.[38] Therefore, considerable attention has recently been paid to the development of emissive drugs.[39,40]

1.4 Objectives of the PhD thesis

Recently, a computational tool named ALTA (Anchor-Based Library Tailoring) was designed in the group of Prof. Caflisch to focus chemical libraries upon docking and prioritizing molecular fragments according to their binding energy. Starting from a relatively small library of 700,000 compounds, 8-(4-hydroxybutyl)-1-methyl-7-phenyl-1H-imidazo[2,1-f]purine-2,4(3H,8H)-dione was identified as a promising EphB4 inhibitor with an inhibition constant of 2.7-2.9 µM and a molecular weight of 353 Da (ALTA_2, Figure 21, left).[41]

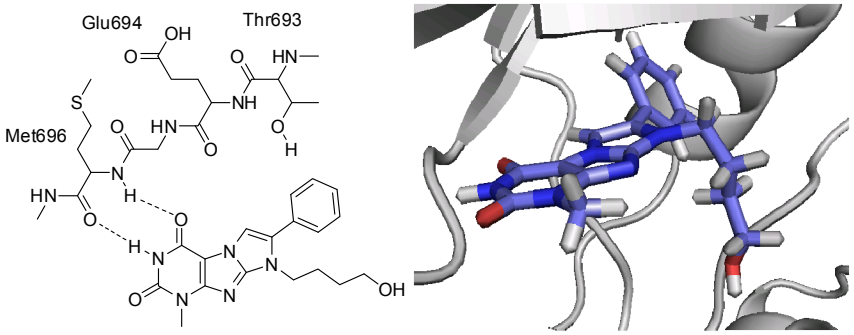


Figure 21. Discovery of ALTA_2

In order to convert this molecule into a clinical candidate, the affinity to the receptor had to be improved. Thus, chapter 2 of this PhD thesis describes experimental studies supported by computational docking that confirmed the expected binding mode, where the phenyl ring of the molecule is buried in a hydrophobic pocket. With these results in hand, lead optimization was performed, resulting in a single-digit nanomolar inhibitor.[18]

However, our lead compound exhibited lower levels of activity in cells, probably due to substrate recognition by efflux transporters. Thus, chapter 3 describes the synthesis of a second generation of inhibitors with improved cellular activities. Their pharmacological properties were assessed in different cell lines and their binding mode was confirmed by X-ray crystallography.

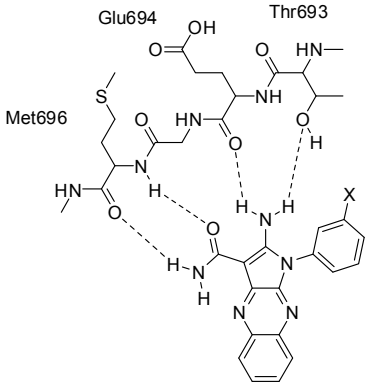


Figure 22. Discovery of a second class of EphB4 inhibitors

In chapter 4, the binding affinity of a second class of EphB4 inhibitors (Figure 22) discovered by automatic docking and subsequent hydrogen bonding penalty calculations[42] was optimized. Their anti-proliferation activity was assessed in the MDA-MB-231 cell line and their fluorescent properties were used for cellular imaging.

1.5 References

(1) Jemal, A.; Bray, F.; Center, M. M.; Ferlay, J.; Ward, E.; Forman, D. Global cancer statistics. *CA. Cancer J. Clin.* **2011**, *61*, 69-90.

(2) Patrick, G. L., *An introduction to medicinal chemistry*. fourth edition ed.; Oxford University Press: 2009.

(3) Rosenberg, B.; VanCamp, L.; Trosko, J. E.; Mansour, V. H. Platinum compounds: a new class of potent antitumour agents. *Nature* **1969**, *222*, 385-386.

(4) Wall, M. E.; Wani, M. C. Camptothecin and taxol: discovery to clinic--thirteenth Bruce F. Cain Memorial Award Lecture. *Cancer Res.* **1995**, *55*, 753-760.

(5) Voet, D.; Voet, J. G., *Biochemistry*. third edition ed.; John Wiley & Sons, New York: 2004.

(6) Traxler, P.; Furet, P. Strategies toward the design of novel and selective protein tyrosine kinase inhibitors. *Pharmacol. Ther.* **1999**, *82*, 195-206.

(7) Cohen, P. Protein kinases--the major drug targets of the twenty-first century? *Nat. Rev. Drug Discov.* **2002**, *1*, 309-315.

(8) Lombardo, L. J.; Lee, F. Y.; Chen, P.; Norris, D.; Barrish, J. C.; Behnia, K.; Castaneda, S.; Cornelius, L. A.; Das, J.; Doweyko, A. M.; Fairchild, C.; Hunt, J. T.; Inigo, I.; Johnston, K.; Kamath, A.; Kan, D.; Klei, H.; Marathe, P.; Pang, S.; Peterson, R.; Pitt, S.; Schieven, G. L.; Schmidt, R. J.; Tokarski, J.; Wen, M. L.; Wityak, J.; Borzilleri, R. M. Discovery of N-(2-chloro-6-methyl- phenyl)-2-(6-(4-(2-hydroxyethyl)- piperazin-1-yl)-2-methylpyrimidin-4- ylamino)thiazole-5-carboxamide (BMS-354825), a dual Src/Abl kinase inhibitor with potent antitumor activity in preclinical assays. *J. Med. Chem.* **2004**, *47*, 6658-6661.

(9) Liu, Y.; Gray, N. S. Rational design of inhibitors that bind to inactive kinase conformations. *Nat. Chem. Biol.* **2006**, *2*, 358-364.

(10) Druker, B. J.; Lydon, N. B. Lessons learned from the development of an abl tyrosine kinase inhibitor for chronic myelogenous leukemia. *J. Clin. Invest.* **2000**, *105*, 3-7.

(11) Pasquale, E. B. Eph receptor signalling casts a wide net on cell behaviour. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 462-475.

(12) Pasquale, E. B. Eph receptors and ephrins in cancer: bidirectional signalling and beyond. *Nat. Rev. Cancer* **2010**, *10*, 165-180.

(13) Noren, N. K.; Pasquale, E. B. Paradoxes of the EphB4 receptor in cancer. *Cancer Res.* **2007**, *67*, 3994-3997.

(14) Kertesz, N.; Krasnoperov, V.; Reddy, R.; Leshanski, L.; Kumar, S. R.; Zozulya, S.; Gill, P. S. The soluble extracellular domain of EphB4 (sEphB4) antagonizes EphB4-EphrinB2 interaction, modulates angiogenesis, and inhibits tumor growth. *Blood* **2006**, *107*, 2330-2338.

(15) Martiny-Baron, G.; Korff, T.; Schaffner, F.; Esser, N.; Eggstein, S.; Marme, D.; Augustin, H. G. Inhibition of tumor growth and angiogenesis by soluble EphB4. *Neoplasia* **2004**, *6*, 248-257.

(16) Martiny-Baron, G.; Holzer, P.; Billy, E.; Schnell, C.; Brueggen, J.; Ferretti, M.; Schmiedeberg, N.; Wood, J. M.; Furet, P.; Imbach, P. The small molecule specific EphB4 kinase inhibitor NVP-BHG712 inhibits VEGF driven angiogenesis. *Angiogenesis* **2010**, *13*, 259-267.

(17) Bardelle, C.; Cross, D.; Davenport, S.; Kettle, J. G.; Ko, E. J.; Leach, A. G.; Mortlock, A.; Read, J.; Roberts, N. J.; Robins, P.; Williams, E. J. Inhibitors of the tyrosine kinase EphB4. Part 1: Structure-based design and optimization of a series of 2,4-bis-anilinopyrimidines. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 2776-2780.

(18) Lafleur, K.; Huang, D.; Zhou, T.; Caflisch, A.; Nevado, C. Structure-based optimization of potent and selective inhibitors of the tyrosine kinase erythropoietin producing human hepatocellular carcinoma receptor B4 (EphB4). *J. Med. Chem.* **2009**, *52*, 6433-6446.

(19) Levine, D. P. Vancomycin: a history. *Clin. Infect. Dis.* **2006**, *42 Suppl 1*, S5-12.

(20) Leach, A. R., *Molecular modelling: principles and applications*. second edition ed.; Pearson Education, London.: 2001.

(21) Goodsell, D. S.; Olson, A. J. Automated docking of substrates to proteins by simulated annealing. *Proteins* **1990**, *8*, 195-202.

(22) Kolb, P.; Caflisch, A. Automatic and efficient decomposition of two-dimensional structures of small molecules for fragment-based high-throughput docking. *J. Med. Chem.* **2006**, *49*, 7384-7392.

(23) Majeux, N.; Scarsi, M.; Caflisch, A. Efficient electrostatic solvation model for protein-fragment docking. *Proteins* **2001**, *42*, 256-268.

(24) Budin, N.; Majeux, N.; Caflisch, A. Fragment-Based flexible ligand docking by evolutionary optimization. *Biol. Chem.* **2001**, *382*, 1365-1372.

(25) Brooks, B. R.; Brooks, C. L., 3rd; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The biomolecular simulation program. *J.Comput. Chem.* **2009**, *30*, 1545-1614.

(26) Schneider, G.; Fechner, U. Computer-based *de novo* design of drug-like molecules. *Nat. Rev. Drug Discov.* **2005**, *4*, 649-663.

(27) Bohm, H. J. LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads. *J. Comput. Aided Mol. Des.* **1992**, *6*, 593-606.

(28) Nishibata, Y.; Itai, A. Automatic Creation of Drug Candidate Structures Based on Receptor Structure - Starting Point for Artificial Lead Generation. *Tetrahedron* **1991**, *47*, 8985-8990.

(29) Jorgensen, W. L.; Ruiz-Caro, J.; Tirado-Rives, J.; Basavapathruni, A.; Anderson, K. S.; Hamilton, A. D. Computer-aided design of non-nucleoside inhibitors of HIV-1 reverse transcriptase. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 663-667.

(30) Vinkers, H. M.; de Jonge, M. R.; Daeyaert, F. F.; Heeres, J.; Koymans, L. M.; van Lenthe, J. H.; Lewi, P. J.; Timmerman, H.; Van Aken, K.; Janssen, P. A. SYNOPSIS: SYNthesize and OPtimize System *in Silico*. *J. Med. Chem.* **2003**, *46*, 2765-2773.

(31) Coupez, B.; Lewis, R. A. Docking and scoring--theoretically easy, practically impossible? *Curr. Med. Chem.* **2006**, *13*, 2995-3003.

(32) Dey, F.; Caflisch, A. Fragment-based *de novo* ligand design by multiobjective evolutionary optimization. *J. Chem. Inf. Model.* **2008**, *48*, 679-690.

(33) Huang, D.; Caflisch, A. Efficient evaluation of binding free energy using continuum electrostatics solvation. *J. Med. Chem.* **2004**, *47*, 5791-5797.

(34) Valeur, B., *Molecular Fluorescence: Principles and Applications*. In Wiley-VCH: Weinheim, Germany: 2002.

(35) Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. C. Green fluorescent protein as a marker for gene expression. *Science* **1994**, *263*, 802-805.

(36) Tsien, R. Y. The green fluorescent protein. *Annu. Rev. Biochem.* **1998**, *67*, 509-544.

(37) Yenugonda, V. M.; Deb, T. B.; Grindrod, S. C.; Dakshanamurthy, S.; Yang, Y.; Paige, M.; Brown, M. L. Fluorescent cyclin-dependent kinase inhibitors block the proliferation of human breast cancer cells. *Bioorg. Med. Chem.* **2011**, *19*, 2714-2725.

(38) Klein, A. V.; Hambley, T. W. Platinum drug distribution in cancer cells and tumors. *Chem. Rev.* **2009**, *109*, 4911-4920.

(39) Kim, D.; Lee, H.; Jun, H.; Hong, S. S.; Hong, S. Fluorescent phosphoinositide 3-kinase inhibitors suitable for monitoring of intracellular distribution. *Bioorg. Med. Chem.* **2011**, *19*, 2508-2516.

(40) Zhao, D.; Wang, W.; Yang, F.; Lan, J.; Yang, L.; Gao, G.; You, J. Copper-catalyzed direct C arylation of heterocycles with aryl bromides: discovery of fluorescent core frameworks. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 3296-3300.

(41) Kolb, P.; Kipouros, C. B.; Huang, D.; Caflisch, A. Structure-based tailoring of compound libraries for high-throughput screening: discovery of novel EphB4 kinase inhibitors. *Proteins* **2008**, *73*, 11-18.

(42) Zhao, H.; Huang, D. Hydrogen bonding penalty upon ligand binding. *PLoS ONE* **2011**, *6*, e19923.

Chapter 2: Structure-Based Optimization of Potent and Selective Inhibitors of the Tyrosine Kinase EphB4

2 Structure-Based Optimization of Potent and Selective Inhibitors of the Tyrosine Kinase EphB4

2.1 Introduction

Angiogenesis, the formation of new blood vessels from pre-existing vessels, has been identified as one of the key steps in human carcinogenesis. In fact, nutrient supply and waste elimination are required for cell proliferation. Because of low toxicity and resistance potential,[1] as well as the possibility to treat a large spectrum of solid tumor types,[2] angiogenesis inhibition is considered a promising target in anticancer therapies. Several studies have implicated erythropoietin-producing human hepatocellular carcinoma receptor (Eph) signaling in sprouting angiogenesis and blood vessel remodeling during vascular development.[3] Furthermore, overexpression of several of the 14 Eph receptors has been linked to tumors and the associated vasculature, suggesting a critical role in tumor-related angiogenesis. In fact, inhibition of binding of EphB4 to its natural ligand ephrinB2 using soluble extracellular domains of EphB4 has been shown to reduce tumor growth in murine tumor xenograft models.[4,5] Thus, inhibition of Eph angiogenic activity has been recognized as an effective strategy for blocking tumor progression and metastasis.

Like all receptor tyrosine kinases, EphB4 is a type-I transmembrane protein. Its extracellular domain is composed of an N-terminal domain necessary for ligand binding, whereas its intracellular domain includes a C-terminal domain and a tyrosine kinase domain. Despite the potential therapeutic importance of EphB4, only four series of small molecule inhibitors are known to date (Figure 1).[6-9] In 2007, Miyazaki and coworkers reported the synthesis of 3-[4-amino-3-(3-chloro-4-fluorophenyl)thieno[3,2-c]pyridin-7-yl]-benzenesulfon-amide (**1**), which is a potent EphB4 inhibitor.[7] One year later, 2,4-bis-anilinopyrimidine derivatives such as **2** showed also high potency as EphB4 inhibitors, and their co-crystallization with human EphB4 highlighted their dual binding mode (Figure 1).[10] The marketed drug dasatinib, with Abl1 and Src as primary targets, also showed a very high affinity to Eph kinases.[11]

High throughput docking is a computational tool frequently used to discover small-molecule inhibitors of enzymes or receptors of known three-dimensional structures.[12,13] Recently, we have developed an efficient computational method (termed ALTA for Anchor-Based Library Tailoring) to focus a chemical library by docking and prioritizing molecular fragments according to their binding energy.[6] From a collection of about 700,000 compounds, ALTA generated a focused library of 21,418 molecules, each containing at least one fragment predicted to bind to the ATP-binding site of EphB4. Automatic

docking of these 21,418 molecules yielded two series of micromolar inhibitors, one of them based on a xanthine scaffold predicted to be involved in two hydrogen bonds with the hinge region that connects the N-terminal and C-terminal lobes of the kinase domain. Further characterization of the commercially available **3** (Figure 1) indicated that this molecule binds to the ATP-binding site, as predicted by the docking calculations.[6] In addition, an analogue with a pendant anisidine chain (**4**, Figure 1) showed similar inhibition properties to **3**, and was active in a cell-based assay.

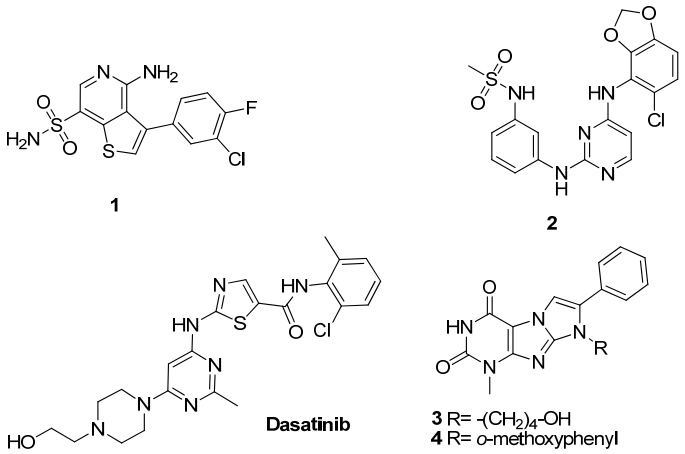


Figure 1. Previously known EphB4 inhibitors.[6-9]

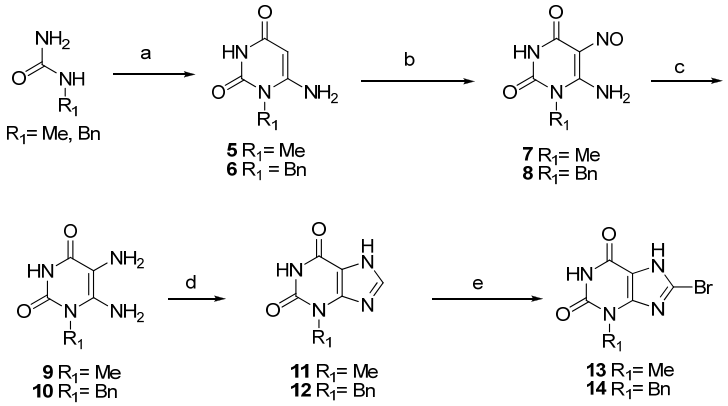
Here, we present a medicinal chemistry campaign aimed at improving the affinity of the micromolar hits **3** and **4** identified by *in silico* screening. The optimization was carried out using the binding mode obtained by automatic docking into EphB4. Chemical synthesis of about 30 derivatives of **3** and **4** yielded six nanomolar inhibitors, one of which (compound **66**) has an IC_{50} of 2 to 5 nM, and shows good selectivity against other protein kinases. The addition of only two heavy atoms (-CH_3 and -OH substituents at the phenyl ring, **4** vs. **66**) has resulted in a ~1,000 times improvement of affinity, which is a remarkable example of the usefulness of structure-based hit modifications.

2.2 Synthesis

Our foreseen modifications to the scaffold of compound **4** ($R_{1\text{-}7}$) for the structure-activity relationship study (SAR) have been summarized at the top of Table 1. Such studies demand a rapid, reliable, and flexible access to a wide variety of potentially active structures with a minimum synthetic variation cost. To succeed in such a goal, a modular synthetic approach was developed. Thus, methyl and benzyl substituted derivatives at R_1 were prepared in parallel as summarized in Scheme 1. The synthesis started by condensation of cyanoacetic acid with commercially available methyl or benzyl urea, to give the corresponding cyanoacetylurea intermediates, which upon treatment with base afforded the desired 1-

alkyl-6-aminouracils (**5**,**6**).[14] Nitrosation at C_5 of the pyrimidine ring with sodium nitrite in acetic acid furnished compounds **7** and **8**, which were subsequently reduced with sodium dithionite to give 1-alkyl-5,6-diaminouracils **9** and **10**.[15] The diamino compounds were immediately refluxed with formic acid to give an amide intermediate, followed by cyclization in basic media, to afford the desired xanthines **11**-**12**. Bromination at C_8 of the xanthine core with Br_2 and sodium acetate in acetic acid led to the formation of the key 8-bromo-xanthines **13**-**14** in excellent overall yield.[16]

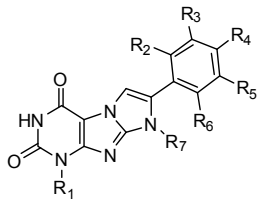
Scheme 1[a]



[a]Reagents and conditions: (a) (i) $NCCH_2CO_2H$, Ac_2O, 60 °C, 1.5 h. ; (ii) NaOH, 90 °C, 30 min. (b) $NaNO_2$, AcOH-H_2O, 25 °C, 12 h. (c) $Na_2S_2O_4$, NH_4OH, 50 °C, 1 h then 25 °C, 8 h. (d) (i) Formic acid, reflux, 3 h. ; (ii) NaOH, reflux, 1 h. (e) Br_2, NaOAc, AcOH, 65 °C, 2 h.

The synthesis of the non-commercially available α-halo-ketones has been summarized in Scheme 2. First, commercially available 1,3-benzodioxole-5-carboxaldehyde (piperonal, **15**) was reacted with methyl magnesium bromide to give secondary alcohol **16**, which was subsequently oxidized in the presence of manganese oxide to give methyl ketone **17**. α-Bromination at the methyl position took place in the presence of phenyltrimethylamoniumtribromide to give alkylating agent **18** in only three steps. 2-Bromo-2´-methyl-3'-hydroxyacetophenone (**24**) and 2-bromo-2´-methyl-4'-hydroxyacetophenone (**25**) were prepared by treatment of the corresponding carboxylic acids with methyllithium to afford methyl ketones **21** and **23**, respectively (Scheme 2). Demethylation of **21** using $AlCl_3$ yielded phenolic derivative **22**. Finally, α-bromination at the methyl group was achieved in the presence of copper(II) bromide in chloroform to give monohalogenated ketones **24** and **25**.

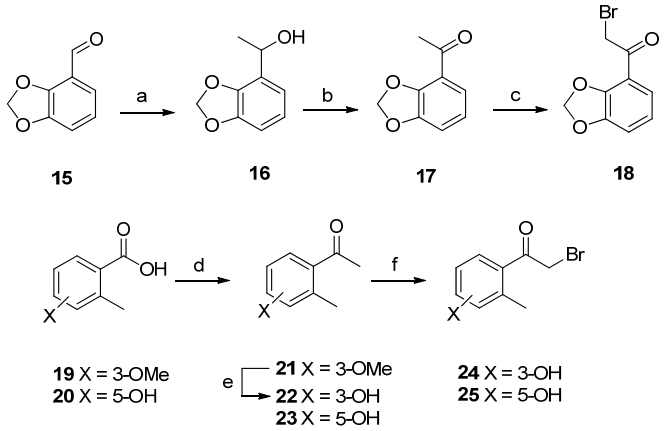
Table 1. EphB4 inhibition data for xanthine derivatives.



Compound	R_1	R_2	R_3	R_4	R_5	R_6	R_7	IC_{50} (nM)[a]
3	Me	H	H	H	H	H	4-hydroxybutyl	7000 (5680)
4	Me	H	H	H	H	H	o-methoxyphenyl	3300 (4350)
45	Bn	H	H	H	H	H	o-methoxyphenyl	> 10000
46	Bn	H	H	F	H	H	o-methoxyphenyl	>10000
69	H	H	H	F	H	H	o-methoxyphenyl	5400
48	Me	H	1,3-dioxole		H	H	o-methoxyphenyl	> 20000
49	Me	H	1,3-dioxole		H	H	4-hydroxybutyl	42% @ 10 µM
50	Me	1,3-dioxole		H	H	H	o-methoxyphenyl	30 % @ 10 µM
51	Me	OMe	H	H	H	H	butyl	> 10000
52	Me	H	OMe	H	H	H	butyl	36% @ 10 µM
53	Me	H	H	OMe	H	H	butyl	> 10000
54	Me	Me	H	H	H	H	o-methoxyphenyl	180 (47)
55	Me	H	Me	H	H	H	o-methoxyphenyl	> 10000
56	Me	H	Me	H	H	H	butyl	37% @ 10 µM
57	Me	H	H	Me	H	H	butyl	38% @ 10 µM
58	Me	OH	H	H	H	H	o-methoxyphenyl	64% @ 10 µM
59	Me	OH	H	H	H	H	butyl	1600
60	Me	H	H	H	OH	H	o-methoxyphenyl	368 (213)
61	Me	H	H	H	OH	H	butyl	691
62	Me	H	H	OH	H	H	o-methoxyphenyl	59% @ 10 µM
63	Me	H	H	OH	H	H	butyl	558
64	Me	Me	OH	H	H	H	o-methoxyphenyl	1200
65	Me	Me	H	OH	H	H	o-methoxyphenyl	236
66	Me	Me	H	H	OH	H	o-methoxyphenyl	5 (1.6)
67	Me	H	H	H	OH	OH	o-methoxyphenyl	5000

[a] IC_{50} values determined at Reaction Biology are given in parentheses. Compounds **3** and **4** were reported previously.[6] Their IC_{50} values were measured again for this study.

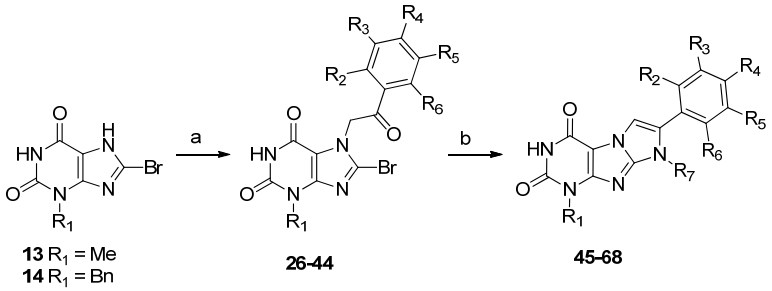
Scheme 2[a]



[a]Reagents and conditions: (a) MeMgBr, THF, -10 °C, 1 h. (b) MnO_2, Et_2O, 25 °C, 48 h. (c) Phenyltrimethylamoniumtribromide, THF, 25 °C, 16 h. (d) MeLi, Et_2O, 25 °C, 3.5 h. (e) $AlCl_3$, PhCl, reflux, 6 h. (f) $CuBr_2$, $CHCl_3$, EtOAc, reflux, 15 h.

Alkylation at the N_7 position of the xanthine core with the appropriate α-haloketone was accomplished using *N,N*-diisopropylethylamine in dimethylformamide providing the 8-bromo-3-alkyl-7-(2-oxo-2-phenyl-ethyl)-xanthine derivatives **26-44** in good yields. Treatment of these intermediates with primary alkyl- or aromatic amines in a sealed tube at 180 °C in ethanol as solvent afforded the desired imidazo[1,2-*f*]xanthine derivatives **45-68** (Scheme 3).

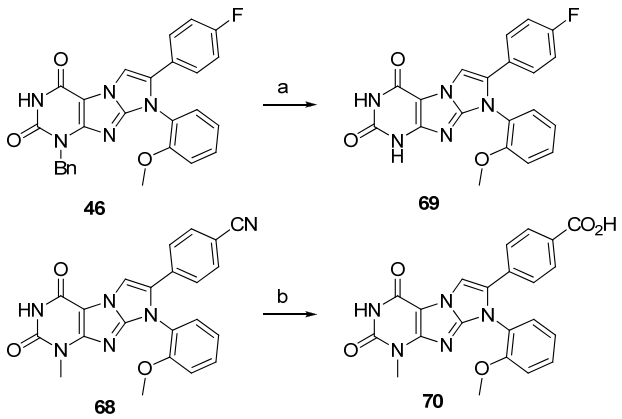
Scheme 3[a]



[a]Reagents and conditions: (a) α-haloketone, DIPEA, DMF, 25 °C, 17 h. (b) Primary amine, EtOH, sealed tube, reflux, 15 h.

Two more derivatives were prepared as chemical probes to address the binding mode of these molecules to EphB4 (Scheme 4). First, the benzyl group in **46** was removed in the presence of Pd/C with ammonium formate to give **69**. Alternatively, the nitrile group in **68** was transformed into the corresponding carboxylic acid (**70**) by hydrolysis with sulfuric acid.

Scheme 4[a]



[a]Reagents and conditions: (a) Pd/C, ammonium formate, MeOH, 140 °C, 3 h. (b) H_2SO_4, H_2O, 120 °C, 2 h.

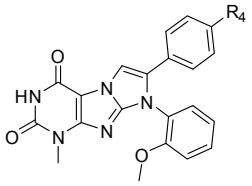
2.3 Results and discussion

The inhibitory activity of the compounds prepared in Schemes 3 and 4 was measured by a fluorescence-resonance energy transfer (FRET)-based enzymatic assay that quantifies inhibition of phosphorylation of a synthetic substrate of EphB4 at K_m concentration of ATP (see Experimental Section).

2.3.1 Experimental validation of the binding mode

The binding mode of compound **3** obtained by automatic docking[6] indicated that the carbonyl C_6=O and amide group N_1-H are involved in hydrogen bonds with the backbone polar groups of Met696 (Figure 2, left). In this model, the phenyl ring of **3** is buried into the hydrophobic pocket, while the 4-hydroxybutyl lateral chain at R_7 points towards the solvent. Visual inspection of such binding mode suggested that another pose could be obtained by a 180° rotation of the first pose (binding mode B, Figure 2, right), so that the phenyl ring points toward the solvent. In pose B, the C_6=O and N_1-H of compound **3** interact with the NH of Met696 and the backbone CO of Glu694, respectively. To discriminate between these two binding modes, we decided to introduce polar substituents at the para position of the phenyl ring. Derivatives with a nitro (**47**), a cyano (**68**), or a highly hydrophilic carboxylic group (**70**) were prepared according to the method described in Schemes 3 and 4. These three derivatives were inactive (Table 2). Furthermore, a hydroxyl substituent at R_4 reduced the affinity by a factor of about 2 to 3 (compare **4** and **62**, Table 1). These results indicate that the phenyl ring of compound **3** more likely fits into the hydrophobic pocket of the ATP-binding site, as suggested by automatic docking (binding mode A in Figure 2).

Table 2. EphB4 inhibition data for compounds synthesized to discriminate between two putative binding modes.



Compound	R4	IC_{50} (nM)
47	NO2	> 20000
68	CN	> 20000
70	CO2H	> 20000

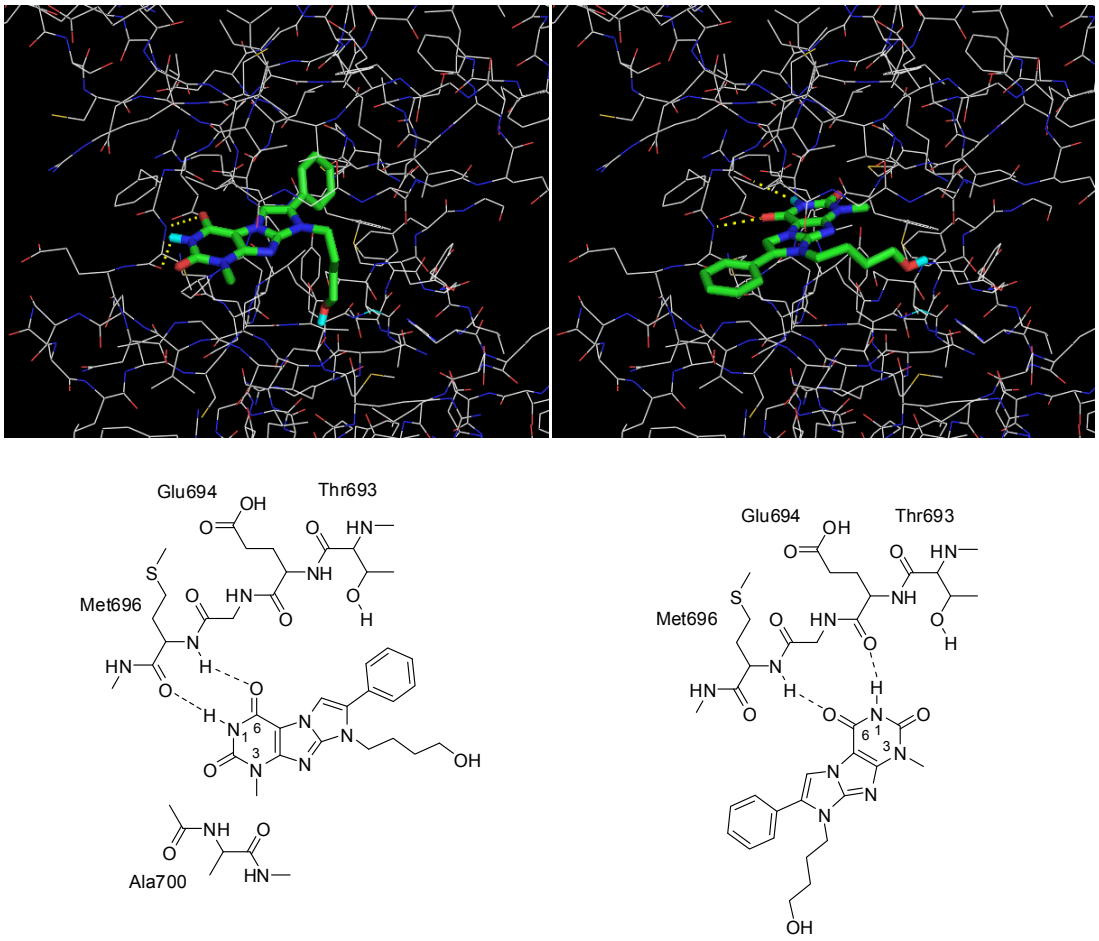


Figure 2. Binding modes of compound **3**. Poses A (left) and B (right) were generated by automatic[6] and manual docking, respectively. (Bottom) The side chain of Phe695 is not shown for clarity.

2.3.2 Lead optimization strategy

Once sufficient evidence about the binding mode had been gathered, we decided to start our optimization campaign from commercially available **4** (compound **5** in Ref. 6). Compounds **3** and **4** differ only at R_7 (alkyl vs. aromatic) and have similar IC_{50} values in the enzymatic assay (Table 1), but only the latter showed activity in a cell-based assay (Table I in Ref. 6). According to the binding mode, the substituent at N_3 of the pyrimidine ring could be involved in additional interactions with Ala700 (Figure 2), so that we decided to start the chemical edition of the molecule at R_1. The methyl group was replaced by a benzyl substituent causing a major loss in the inhibitory activity of these molecules (**45** and **46** vs. **4**). Compound **69**, with a N_3-H bond was not more potent than **4**. Since no improvement was achieved by modification of the pyrimidine ring, we decided to focus our efforts on a different region of the molecule.

The ATP-binding site of protein kinases can be divided into five different sub-pockets.[17] The size of the hydrophobic pocket is controlled by a so called gatekeeper residue, and it is well known that not only affinity but also selectivity can be improved by fully exploring this site.[17] In fact, only around 20% of the 518 human kinases have a small gatekeeper residue. Since EphB4 belongs to this class having a threonine gatekeeper (Thr693), we envisioned that a straightforward strategy to improve affinity would stem from modifications of the substitution pattern at the phenyl ring. In sharp contrast to compound **2** and other previously developed inhibitors of EphB4,[8,10] a dioxole ring in relative positions R_3, R_4 (**48**-**49**) or R_2, R_3 (**50**) of the aromatic ring significantly reduced the activity. A methoxy group with different substitution patterns (**51**, **52**, **53**) also suppressed the inhibitory activity of the corresponding molecules. According to binding mode A (Figure 2) the space around the phenyl ring is rather limited, thus restricting the size of the substituents that might improve steric complementarity. Furthermore, the phenyl ring is close to the carbonyl group of Glu664, the hydroxyl and carbonyl groups of Ser757, and the NH group of Asp758, which can act as hydrogen bonding partners. Based on these observations, we decided to examine less sterically demanding substituents such as methyl or hydroxyl groups, which are expected to fulfill the nearby hydrogen bonding capacity.

Notably, a methyl substituent at position R_2 of the benzene ring (**54**) showed an IC_{50} value close to 100 nM. In contrast, R_3 and R_4 methyl substituted derivatives (**55-57**) were almost inactive. Furthermore, introduction of a hydroxyl group at R_5 (**60**) was also beneficial with an IC_{50} of about 200-400 nM. Some inhibition activity was also detected for compounds bearing the OH group at R_2 (**58**) and R_4 (**62**) of the phenyl ring. In these cases, replacing the anisidine lateral chain at R_7 (**58**, **60**, and **62**) for an alkylic chain such as butyl (**59**, **61**, and **63**) seemed to have only a limited influence in the inhibitory activity of these

molecules as previously observed for the commercially available compounds with a propyl or a butyl chain at R_7.[6] With these results in hand, we decided to explore how the combination of the most favored substitution patterns could influence the inhibitory activity. Thus, the methyl substituent was kept at R_2 and a hydroxyl group was added at the relative positions R_3 (**64**), R_4 (**65**), and R_5 (**66**) of the benzene ring. Strikingly, the combination of a methyl and a hydroxyl group at R_2 and R_5 respectively, yielded compound **66** which has a ~1,000 fold higher affinity compared to the original hits obtained by docking, i.e., compounds **3** and **4**.

As mentioned above, the addition of only one heavy atom, CH_3 at R_2, resulted in a factor of 20-100 improvement (**54** (180 nM) vs. **4** (3300 nM), or **66** (5 nM) vs. **60** (691 nM)). This observation led us to further investigate the role of the methyl group, in particular if it stabilizes the orientation of the phenyl ring required for binding. Conformational analysis was performed on **4**, **54**, **60** and **66** by exhaustive sampling of the dihedral angles involved in the rotation of the phenyl and o-methoxyphenyl rings (γ1 and γ2 in Figure 3), followed by geometry optimization of the resulting structures using semi-empirical methods (AM1 Hamiltonian) and density functional theory (B3LYP level). For each of the four inhibitors, the local minima are distributed into four sets of conformers, which are separated by rotation barriers. The docked conformation lies in one of these four basins, with similar energy values (maximal difference in energy of 0.08 kcal/mol for **4**, 0.22 kcal/mol for **54**, 0.26 kcal/mol for **60**, and 0.14 kcal/mol for **66**), which suggests a quasi-equal distribution of the population in each of the four conformers. Notably, the presence of a methyl group at R_2 seems to restrict the accessible conformations more than a hydroxyl group at R_5 (compare the plots obtained for **54** and **66**, and for **4** and **60** in Figure 3). In addition, the conformational strain, which is the energy difference between the minimized bound conformation and the lowest energy conformation of the isolated ligand, was evaluated for each molecule. The similar values obtained for the strain energy of **54** (0.5 kcal/mol) and **4** (0 kcal/mol), as well as **66** (0.7 kcal/mol) and **60** (0.3 kcal/mol), indicate that the methyl group contributes to a gain in intermolecular van der Waals energy rather than in strain energy (see also the subsection **Binding mode of compound 66 investigated by MD simulations**).

During the preparation of this manuscript we discovered in the literature a series of inhibitors of the tyrosine kinase Lck with a 2,4-dianilinopyrimidine scaffold that have a very similar SAR for the phenyl substituents to the one observed for compounds **4**, **54**, **60** and **66**.[18] Note that the phenyl substituent in both series of compounds is located in the hydrophobic pocket, but it is connected to the 2-anilinopyrimidine core by a -NH- linker in the Lck inhibitors, whereas the phenyl ring is attached directly to the xanthine scaffold in the EphB4 inhibitors described here.

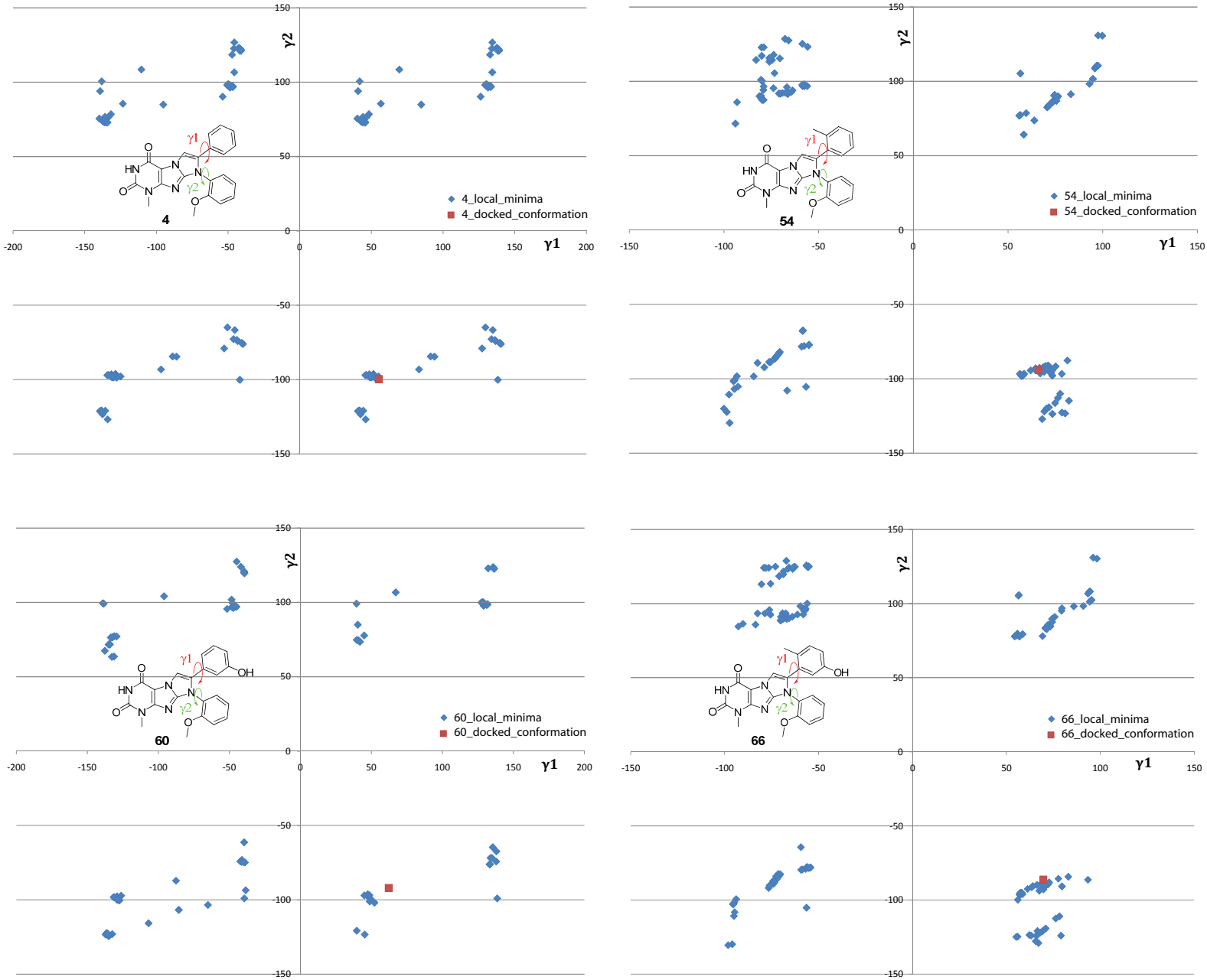


Figure 3. Conformational analysis on **4**, **54**, **60**, **66**. The local minima obtained by ab initio minimization of each rotameric state are plotted against the dihedral angles γ1 and γ2 (blue). The global minima are indicated in red.

2.3.3 Selectivity profile

To assess the specificity of kinase inhibitors, it is useful to distinguish between local and global selectivity profiles, which reflect the inhibitory activity of the tested compound on a single branch of the kinome dendrogram and on the whole kinome, respectively. The local selectivity of compound **66** was tested against a panel of 11 Eph receptor kinases by an enzymatic assay with [γ-^{33}P]ATP (Reaction Biology Corporation). The IC_{50} values measured for this compound against 10 of the 11 Ephs are in the low nM range (1-40 nM) while an IC_{50} value of 1.1 μM is observed for EphA7 (Table 3). These IC_{50} values are consistent with the very high sequence identity (60%-90%) of Ephs, and the bulkier gatekeeper residue in EphA7 (isoleucine) with respect to the other Eph kinases (threonine). To evaluate the global selectivity, enzymatic assays (at single concentration of inhibitor) were performed for compounds **66** and **54** using a panel of 85 kinases (National Centre for Protein Kinase Profiling at the University of Dundee, Figure 4). Out of these 85 kinases, only five (EphA2, EphB3, Src, Lck, and Yes1) and three (EphA2, Lck, Yes1) are very strongly inhibited by compounds **66** and **54**, respectively (less than 10% activity remaining compared to a DMSO control at concentrations of 1 μM of **66** and 3 μM of **54**, see section 2.6). It is important to notice that these five kinases have a threonine as gatekeeper, as well as CSK, BTK and HER-4, which show relatively strong inhibition by **66** (between 10 and 50% activity remaining compared to a DMSO control).

Table 3. Local selectivity of compound **66**. These IC_{50} values were measured at Reaction Biology Corporation.

Kinases	IC_{50} (nM)
EphA1	2.9
EphA2	2.3
EphA3	40
EphA4	3.3
EphA5	3.0
EphA7	1118
EphA8	4.5
EphB1	1.1
EphB2	1.2
EphB3	15
EphB4	1.6

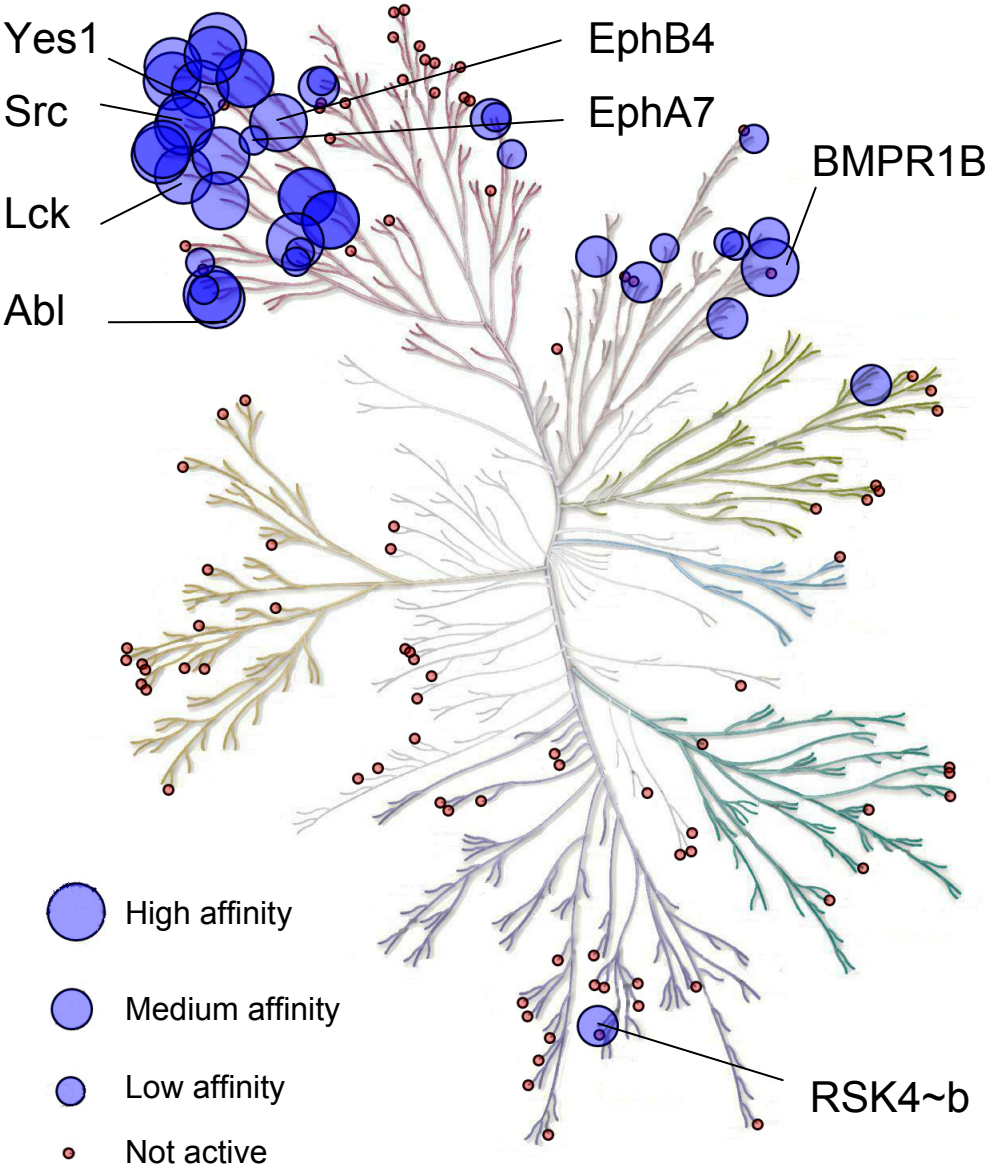


Figure 4. Selectivity profile of compound **66**. The circles correspond to all kinases tested: inhibition of activity was measured in enzymatic assays with radiolabeled ATP at Reaction Biology Corporation and University of Dundee for 11 and 85 kinases, respectively, while binding affinity of 50 kinases was measured at Ambit Biosciences Corporation. Enzymatic assays: high, medium, low, and no compound affinity for kinase activity (with respect to DMSO control) of <10%, 10%-30%, 30%-60%, and >60%, respectively. Binding assay: high, medium, low, and no compound affinity for kinase activity (with respect to DMSO control) of <1%, 1%-10%, 10%-30%, and >30%, respectively. The dendrogram is adapted from a previously reported picture.[19]

The local and global selectivity profiles prompted us to test the binding of compound **66** against a set of 49 kinases (Ambit Biosciences Corporation) selected among the nearly 100 that are predicted to have a small gatekeeper based on sequence and structure analysis.[20] Interestingly, in competition binding

assays, compound **66** shows significant affinity (kinase activity < 10% with respect to DMSO control) to only 15 of these 49 kinases (see section 2.6).

Considering all these selectivity tests (Figure 4), compound **66** has been profiled against a total of 143 kinases (three kinases, EphA2, EphB3 and EphB4 were tested twice) about half of which were chosen because of their small gatekeeper residue. Only 21 kinases with a threonine gatekeeper showed high inhibition by **66**, and 10 of them are Eph kinases (Figure 4). Among the kinases with a gatekeeper different from threonine, only GCK and MLK1 (both with methionine as a gatekeeper) are inhibited by compound **66**. Interestingly, the selectivity of **66** is consistent with the recently published kinase tree constructed by using SAR data,[21] as the kinases that bind compound **66** colocalize in a small branch of the SAR-based dendrogram (Abl, Src and Eph families, see section2.6). Moreover, EphA7 is located in another branch of the SAR-based dendrogram in agreement with the 100 to 1000 lower affinity with respect to the other 10 Eph kinases.

The profiles provide evidence for a good specificity of compound **66**, as it shows significant inhibition for a relatively small fraction of the human kinome, which is comparable, or even more specific, than some type I kinase inhibitors currently used as drugs, e.g., dasatinib[9] and sunitinib.[22] The high selectivity may originate from fully exploring the inner hydrophobic pocket whose size is controlled by the gatekeeper. In addition, the rigidity of this compound may contribute to its selectivity.

2.3.4 Binding mode of compound 66 investigated by MD simulations

Although the SAR of the 28 xanthine derivatives (Tables 1 and 2) is consistent with the pose of compounds **3** and **4** predicted by automatic docking,[6] explicit water molecular dynamics (MD) simulations were carried out to further validate the binding mode of compound **66**, which is the most potent of all derivatives. The MD run started from the pose of **66** obtained by automatic flexible-ligand docking (see computational methods). The structure of EphB4 is stable in the 45-ns simulation as indicated by the time series of the Cα RMSD (Figure 5). Importantly, the hydrogen bonds with the hinge region (i.e., between the backbone polar groups of Met696 and the pyrimidine ring of compound **66**) are conserved during the entire run, except for some transient ruptures of the one involving the carbonyl group of Met696 (Figure 6a,b). Furthermore, the side chain of Ser757 is involved in a rather stable hydrogen bond with the methoxy oxygen of compound **66** (Figure 6c). On the other hand, a change in the orientation of the hydroxyl group of compound **66** was observed after about 12 ns, probably due to the absence of Mg in the simulation. The two hydrogen bonds with the side chain of Glu664 and the backbone NH of Asp758, broke apart with concomitant formation of a rather stable hydrogen bond with the backbone CO of Ser757 (Figure 6d). This change in the hydrogen bonding pattern of the hydroxyl

group occurred immediately after the rotation of the peptide bond between Ser757 and Asp758 (Figure 6) and persisted until the end of the MD run. In other words, the reorientation of the Ser757 CO and Asp758 NH groups promoted the new hydrogen bond pattern.

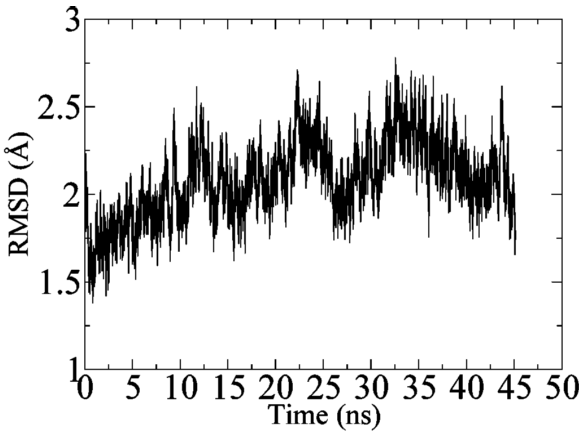


Figure 5. Time series of the Cα root mean square deviation (RMSD) of residues 613-883 of the kinase domain of EphB4 with respect to the X-ray structure (PDB code 2VWX). The loop 772-778 was not taken into account because it is not present in the crystal structure.

Overall, the 45-ns MD simulation of the complex of EphB4 with compound **66** shows that the major features of the binding mode predicted by docking are stable. These include the orientation of the phenyl in the hydrophobic pocket and the two hydrogen bonds with the hinge region. It is interesting to note that **66** forms by its hydroxyl substituent one or two additional hydrogen bonds (1.5 to 3 kcal/mol, corresponding to a factor of 10 to 100 in binding affinity) compared to compound **54**, which is consistent with its ~30-35 higher affinity than the one of **54** (Table 1). Furthermore, the methyl group at R_2 is involved in favorable van der Waals interactions with the side chains of Val629, Ala645, and Thr693 throughout the simulation (see section 2.6). These van der Waals interactions contribute about -1.5 kcal/mol to the binding energy (corresponding to a factor of 10 in binding affinity). Compounds **66** and **60** differ only by this methyl group, so that the former has a more favorable desolvation energy upon binding, which, taken together with the additional van der Waals interactions with EphB4, explains the higher potency of compound **66** vs. **60** (Table 1).

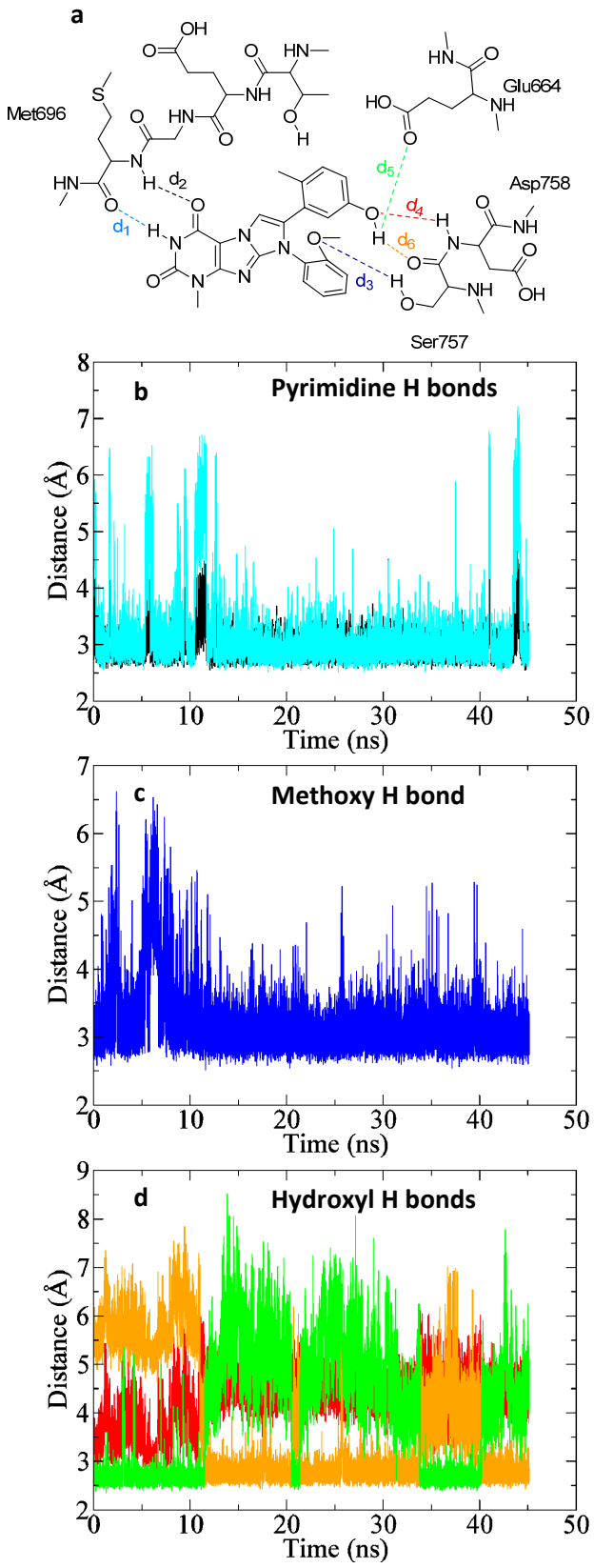


Figure 6. Hydrogen bonds between compound **66** and the ATP-binding site of EphB4. (a) Schematic illustration. (b-d) Time series of hydrogen-bond distances, i.e., the distance between the donor and acceptor atoms. The colors are consistent with those used in part (a).

2.4 Conclusions

We had previously discovered a series of micromolar inhibitors of the tyrosine kinase EphB4 by fragment-based high-throughput docking of a library of about 700,000 compounds.[6] Here, we describe the structure-guided improvement of these relatively weak inhibitors by chemical synthesis that has culminated into the discovery of a single-digit nanomolar inhibitor **66**, with a remarkable increase in ligand efficiency (ratio of the Gibbs free energy to the number of non-hydrogen atoms)[23] from 0.26 to 0.37 kcal/mol per non-hydrogen atom (for inhibitors **4** and **66**, respectively). Noticeably, compound **66** has only two non-hydrogen atoms more than **4** (a methyl group and a hydroxyl group on the same benzene ring), and these small but very effective substituents were suggested by analysis of the binding mode obtained by automatic docking. Furthermore, the predicted binding mode of compound **66** into the ATP-binding site of EphB4 is stable in a 45-ns MD simulation with explicit solvent. Analysis of the MD trajectory suggests that the higher potency of **66** with respect to **4** is due to favorable van der Waals interactions of the methyl group with the side chains of Val629, Ala645, and the gatekeeper residue (Thr693), as well as a hydrogen bond between the hydroxyl group and the backbone CO of Ser757. In addition, quantum mechanics calculations indicate that the beneficial effect of the methyl group is not due to a lower conformational strain upon binding, but rather to the van der Waals interactions characterized by MD simulation. Although the SAR of all the compounds synthesized in this work is consistent with the binding mode of compound **66** obtained by docking and slightly refined by explicit solvent MD, its definitive validation awaits for X-ray crystallography, which is being pursued vigorously and will be reported in due course. Finally, compound **66** shows a relatively good selectivity profile which is an important requirement for further consideration as a lead compound.

2.5 Experimental Section

2.5.1 Docking

The flexible ligand docking employed here for compounds **3** and **66** is a fragment-based approach as it exploits the optimal binding modes of mainly rigid fragments. The approach consists of four consecutive steps that have been detailed in a previous work.[24] Briefly, the four steps are: (1) Decomposition of the ligand into mainly rigid fragments by the program DAIM,[25] (2) fragment docking with evaluation of electrostatic solvation[26,27] by the program SEED,[28,29] (3) flexible docking using the position and orientation of its fragments as anchors by the program FFLD,[30,31] and (4) final minimization by CHARMM.[32,33] The protein structure (PDB file 2VWX) was kept rigid in all steps.

2.5.2 Parametrization of compound 66 and MD simulations

Initial Mulliken partial charges of 1,7,8-trimethyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione were obtained by optimization with Gaussian03 at the HF/6-31G(d) level.[34] Partial charges of hydrogen atoms in the methyl groups were set to 0.09 e, and the excessive positive charge was added to the adjacent carbon atom. To determine the partial charges on the remaining atoms, interaction energies between each polar group and a TIP3 water molecule[35] were evaluated as the difference between the supermolecule energy and the sum of the individual monomer energies. The partial charges on the polar groups were then manually adjusted to fit the interaction energies calculated by CHARMM to the *ab initio* values. Bond lengths and angle values were also modified to reproduce the quantum mechanically obtained geometries. Normal mode contributions generated with the MOLVIB program were adjusted to scaled HF/6-31G(d) values. Final steps included introduction of the previously parametrized phenol and methyl moieties,[34] followed by optimization of the remaining missing parameters.

The coordinates of EphB4 were downloaded from the PDB database (2VWX). As the X-ray structure contains only residues 608 to 888, the -$COCH_3$ group and -$NHCH_3$ group were added to the N-terminal Lys608 and C-terminal Ala888, respectively. To reproduce neutral pH conditions the side chains of aspartates and glutamates were negatively charged, those of lysines and arginines positively charged, and histidines were considered neutral. The protein was immersed in an orthorhombic box of pre-equilibrated water molecules. The size of the box was chosen to have a minimal distance of 13 Å between the boundary and any atom of the protein. The program VMD[36] was used for setting up the simulation system, while minimization, heating, and production runs were performed with NAMD[37] using the CHARMM22 force field[34] and the TIP3P model of water. Periodic boundary conditions were applied and the particle-mesh Ewald approach[38] was used for the long-range electrostatics. The van der Waals interactions were truncated at a cutoff of 12 Å and a switch function is applied starting at 10 Å. The MD simulation was performed at constant temperature of 310 K by applying the Langevin thermostat and constant pressure of 1 atm using the Nose-Hoover Langevin piston pressure control.[39,40] The time step was 2 fs and the SHAKE algorithm[41] was used to fix the length of covalent bonds involving hydrogen atoms. Coordinate sets were saved every 2 ps.

2.5.3 Evaluation of strain energy by quantum mechanics

Systematic rotation of rotatable bonds in increments of 10° and 20° for the two phenyls and the methoxy, respectively, was performed using CHARMM, and the resulting conformations were minimized with Gaussian03 using the AM1 Hamiltonian. Removal of duplicate conformations was done using the Ultrafast Shape Recognition protocol[42] supplemented by a chirality descriptor that distinguishes

molecules that are mirror images of each other. Geometry optimization of the resulting structures was done at the B3LYP level using the PC GAMESS version of the GAMESS (US) quantum chemical package.[43] Redundant conformations were discarded by a second Shape Recognition analysis. The strain energy was calculated as the energy difference between the minimized bound conformation and the global minimum obtained by the systematic search.

2.5.4 Chemistry

All reactions, unless otherwise stated, were carried out under a nitrogen atmosphere using standard Schlenk-techniques. All reagents were used as received unless otherwise noted. Solvents were purchased in the best quality available, degassed by purging thoroughly with nitrogen and dried over activated molecular sieves of appropriate size. Alternatively, they were purged with argon and passed through alumina columns in a solvent purification system (Innovative Technology). Reactions were monitored by thin layer chromatography (TLC) using Merck TLC silica gel 60 F_{254}. Flash column chromatography was performed over silica gel (230-400 mesh). NMR spectra were recorded on AV2 400 or AV2 500 MHz Bruker spectrometers. Chemical shifts are given in ppm. The spectra are calibrated to the residual ^{1}H and ^{13}C signals of the solvents. Multiplicities are abbreviated as follows: singlet (s), doublet (d), triplet (t), quartet (q), doublet-doublet (dd), quintet (quint), septet (sept), multiplet (m), and broad (br). Melting points were determined on a Büchi Melting Point B-540 instrument. Methylurea, benzyl urea, benzo[*d*][1,3]dioxole-4-carbaldehyde (piperonal **15**), 2-methyl-3-methoxybenzoic acid (**19**) and 2-methyl-5-hydroxybenzoic acid (**20**) were purchased from Fluka. The following compounds: 2-bromo-1-phenyl-ethanone, 2-bromo-1-(4-fluorophenyl)-ethanone, 2-bromo-1-(4-methoxyphenyl)-ethanone, 2-bromo-1-(3-methoxyphenyl)-ethanone, 2-bromo-1-(2-methoxyphenyl)-ethanone, 2-chloro-1-(4-hydroxyphenyl)-ethanone, 2-bromo-1-(3-hydroxy-phenyl)-ethanone, 2-bromo-1-(2-hydroxy-phenyl)-ethanone, 2-chloro-1-(3,4-dihydroxy-phenyl)-ethanone, 2-bromo-1-(4-nitro-phenyl)-ethanone, 2-bromo-4′-cyanoacetophenone, 2-chloro-1-(4-hydroxy-2-methylphenyl)-ethanone, 2-bromo-1-(4-methylphenyl)-1-ethanone were purchased from Fluka. 2-Bromo-1-(3-methylphenyl)-1-ethanone and 2-bromo-1-(2-methylphenyl)-1-ethanone were purchased from Synchem and 1-(1,3-benzodioxol-5-yl)-2-bromo-1-ethanone was purchased from Acros Organics.

High-resolution electrospray ionization mass spectrometry was performed on a Finnigan MAT 900 (Thermo Finnigan, San Jose, CA, USA) double-focusing magnetic sector mass spectrometer. Ten spectra were acquired. A mass accuracy ≤2 ppm was obtained in the peak matching acquisition mode by using a solution containing 2 µL PEG200, 2 µL PPG450, and 1.5 mg NaOAc (all obtained from Sigma-Aldrich, Buchs, Switzerland) dissolved in 100 mL MeOH (HPLC Supra grade, Scharlau, E-Barcelona) as internal

standard. The purity of all tested compounds was determined by HPLC on a Finnigan Voyager GC8000 Top spectrometer using an Interchim Strategy (2.2 µm, 2.1×100 mm) column by running a 5-80% gradient for water (+ 0.1% HCOOH + 0.01% TFA)/ CH_3CN. Flow rate was 200 µL/min, and UV detection was set to 254 nm. Unless otherwise stated, all the compounds showed ≥95% purity.

1-(Benzo[*d*][1,3]dioxol-4-yl)ethanol (16)[44]

A solution of MeMgBr in THF (1 M, 9.9 mL, 9.9 mmol, 1.5 eq) was diluted with THF (10mL) and cooled to -10 °C. A solution of benzo[*d*][1,3]dioxole-4-carbaldehyde **15** (1.0 g, 6.6 mmol) in 9.9 mL THF was slowly added, and the reaction was stirred for 1 h. The reaction mixture was then quenched by pouring it into 100 mL of ice cold sat. ammonium chloride and the aqueous layer was then extracted with Et_2O. The organic extracts were dried over $MgSO_4$ and evaporated under reduced pressure to provide the desired compound as a colorless oil (971 mg, 88% yield). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 6.87 (ddd, *J* = 7.8, 1.4, 0.4 Hz, 1H), 6.82 (dd, *J* = 7.8, 7.5 Hz, 1H), 6.75 (dd, *J* = 7.5, 1.4 Hz, 1H), 5.96 (d, *J* = 1.4 Hz, 1H), 5.95 (d, *J* = 1.4 Hz, 1H), 4.99 (q, *J* = 6.5 Hz, 1H), 1.52 (d, *J* = 6.5 Hz, 3H), OH not observed; ^{13}C NMR (100 MHz, $CDCl_3$): δ = 147.3, 143.9, 127.3, 121.7, 118.7, 107.7, 100.8, 66.2, 23.3; IR (film): $\tilde{\upsilon}$ = 3359, 2972, 2887, 1455, 1245, 1039, 727 cm^{-1}; MS (ESI): *m/z*: calcd for $C_9H_{10}NaO_3$: 189.1, found: 189.0 $[M + Na]^+$.

1-(Benzo[*d*][1,3]dioxol-4-yl)ethanone (17)[44]

A mixture of 1-(Benzo[*d*][1,3]dioxol-4-yl)ethanol **16** (971 mg, 5.84 mmol) and MnO_2 (5.08 g, 58.4 mmol) in Et_2O (36 mL) was stirred vigorously for 48 h. The reaction mixture was then filtered through a path of celite, and evaporated under reduced pressure to give the desired compound as a white solid (679 mg, 70% yield). mp 95-97 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.37 (dd, *J* = 8.2, 1.3 Hz, 1H), 6.97 (dd, *J* = 7.6, 1.3 Hz, 1H), 6.87 (dd, *J* = 8.2, 7.6 Hz, 1H), 6.09 (s, 2H), 2.60 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 195.4, 148.6, 147.9, 121.4, 121.2, 120.3, 112.5, 101.5, 30.2; IR (film): $\tilde{\upsilon}$ = 3068, 2908, 1671, 1625, 1450, 1361, 1283, 1240, 1179, 1060, 1022, 950, 727 cm^{-1}; MS (EI): *m/z*: calcd for $C_9H_8O_3$: 164.1, found: 163.9 $[M]^+$.

1-(Benzo[*d*][1,3]dioxol-4-yl)-2-bromoethanone (18)[44]

To a solution of 1-(Benzo[*d*][1,3]dioxol-4-yl)ethanone **17** (679 mg, 4.14 mmol) in anhydrous THF (19 mL) at 0 °C was added phenyltrimethylammonium tribromide (1.55 g, 4.14 mmol). The reaction mixture was stirred at room temperature for 16 h, concentrated, and re-dissolved in EtOAc. The organic layer was washed with water (2 × 50 mL) and brine (1 × 50 mL), dried over $MgSO_4$, filtered, and evaporated under reduced pressure. Purification by column chromatography on silica gel (toluene/hexane 1:4), followed by recristallization in hexane/EtOAc afforded the desired compound as a white solid (309 mg, 30% yield). mp 75-77 °C; ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.42 (dd, *J* = 8.2, 1.2 Hz, 1H), 7.01 (dd, *J* = 7.6, 1.2 Hz, 1H), 6.91 (dd, *J* = 8.2, 7.6 Hz, 1H), 6.11 (s, 2H), 4.48 (s, 2H); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 188.2, 148.5,

147.8, 121.9, 121.7, 117.1, 113.3, 101.8, 34.9; IR (film): $\tilde{\upsilon}$ = 3002, 2942, 2917, 1693, 1624, 1451, 1387, 1248, 1204, 1177, 1048, 942, 768, 723 cm^{-1}; MS (ESI): *m/z*: calcd for $C_9H_7BrNaO_3$: 264.9, found: 265.0 $[M + Na]^+$.

1-(3-Methoxy-2-methylphenyl)-ethanone (21)

To a solution of 2-methyl-3-methoxybenzoic acid **19** (1.5 g, 9.02 mmol) in dry diethyl ether (68 mL) at 0 °C, methyllithium (1.6 M in Et_2O, 22 mL, 36.1 mmol) was added dropwise with vigorous stirring. The reaction was allowed to warm to room temperature and stirred for an additional 3.5 h. The reaction was quenched by pouring into ice cold 10% HCl (50 mL) and the aqueous layer was then extracted with Et_2O (3×50 mL). The organic solution was washed with brine, dried over $MgSO_4$, filtered, and concentrated under reduced pressure. Purification by column chromatography on silica gel (hexane/EtOAc 16:1) afforded the desired compound as a colorless oil (613 mg, 41% yield). ^1H NMR (500 MHz, $CDCl_3$): δ = 7.22 (dd, *J* = 8.1, 7.6 Hz, 1H), 7.15 (dd, *J* = 7.6, 1.1 Hz, 1H), 6.95 (d, *J* = 8.1 Hz, 1H), 3.85 (s, 3H), 2.55 (s, 3H), 2.32 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 203.1, 158.2, 140.6, 126.1, 125.9, 120.0, 112.7, 55.8, 30.4, 12.6; IR (film): $\tilde{\upsilon}$ = 3001, 2960, 2938, 2836, 1684, 1577, 1456, 1258, 1051, 781 cm^{-1}; MS (ESI): *m/z*: calcd for $C_{10}H_{12}NaO_2$: 187.1, found: 187.0 $[M + Na]^+$.

1-(3-Hydroxy-2-methylphenyl)-ethanone (22)

A solution of 1-(3-methoxy-2-methylphenyl)-ethanone **21** (450 mg, 2.74 mmol) in chlorobenzene (11 mL) was treated with $AlCl_3$ (475 mg, 3.56 mmol) at room temperature and then refluxed for 5 h. A second portion of $AlCl_3$ (475 mg, 3.56 mmol) was added, and the resulting mixture was refluxed for an additional hour, till no starting material was detected by TLC. The reaction mixture was carefully treated with 1 M HCl at 0 °C. The aqueous phase was extracted with EtOAc, and the combined organic layers were dried with $MgSO_4$. Evaporation of the solvent and column chromatography on silica gel (hexane/EtOAc 6:1) afforded the desired compound as a yellow solid (328 mg, 79% yield). mp 122-124 °C; ^1H NMR (400 MHz, $CDCl_3$): δ = 7.18 (dd, *J* = 7.8, 1.2 Hz, 1H), 7.12 (t, *J* = 7.8 Hz, 1H), 6.93 (dd, *J* = 7.8, 1.2 Hz, 1H), 5.34 (s, 1H), 2.57 (s, 3H), 2.36 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 203.4, 154.7, 140.4, 126.2, 123.3, 120.9, 118.0, 30.1, 12.4; IR (film): $\tilde{\upsilon}$ = 3221, 1652, 1577, 1466, 1360, 1277, 782 cm^{-1}; MS (ESI): *m/z*: calcd for $C_9H_{10}NaO_2$: 173.1, found: 173.0 $[M + Na]^+$.

1-(5-Hydroxy-2-methylphenyl)-ethanone (23)

To a solution of 2-methyl-5-hydroxybenzoic acid **20** (900 mg, 5.9 mmol) in THF (6.5 mL) at -78 °C was added MeLi (1.6 M in ether, 11 mL, 17.7 mmol). The reaction was stirred 17 h while slowly warming up to room temperature. EtOAc (10 mL) was slowly added to the reaction mixture, which was subsequently washed (10 mL of 1 M HCl, 10 mL of water, and 10 mL of brine). After drying the organic layer over

$MgSO_4$ and evaporation of the solvent under reduced pressure, the residue was purified by flash chromatography on silica gel (EtOAc/hexane 1:10) to afford desired compound as a white solid (608 mg, 68% yield). mp 129-131 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.17 (d, *J* = 2.6 Hz, 1H), 7.10 (d, *J* = 8.2 Hz, 1H), 6.88 (dd, *J* = 8.2, 2.6 Hz, 1H), 5.32-5.37 (m, 1H), 2.55 (s, 3H), 2.43 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 201.9, 153.4, 138.5, 133.1, 130.2, 118.7, 116.1, 29.5, 20.6; IR (film): $\tilde{\upsilon}$ = 3188, 2929, 1657, 1603, 1456, 1434, 1302, 1221, 734 cm^{-1}; MS (ESI): *m/z*: calcd for $C_9H_{10}NaO_2$: 173.1, found: 173.0 $[M + Na]^+$.

General procedure for the α-bromination of acetophenones

A solution of acetophenone (1 eq) in $CHCl_3$ (0.18 M) was added to a refluxing solution of copper (II) bromide (1.99 eq) in EtOAc (0.45 M). The mixture was then refluxed for 10 h. The solution was filtered through celite and concentrated under reduced pressure to afford a green solid. The solid was purified by flash chromatography on silica gel (toluene) to afford the corresponding products in pure form. This method was used to obtain **24** and **25**.

2-Bromo-1-(3-hydroxy-2-methylphenyl)-ethanone (24)

White solid; Yield: 61%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.15-7.17 (m, 2H), 6.95-6.98 (m, 1H), 5.14 (s, 1H), 4.39 (s, 2H), 2.34 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 195.1, 154.7, 137.2, 126.5, 124.4, 120.6, 118.8, 34.2, 12.4; IR (film): $\tilde{\upsilon}$ = 3420, 2933, 1683, 1581, 1466, 1327, 796 cm^{-1}; MS (ESI): *m/z*: calcd for $C_9H_9BrNaO_2$: 250.9, found: 251.0 $[M + Na]^+$.

2-Bromo-1-(5-hydroxy-2-methylphenyl)-ethanone (25)

White solid; Yield: 50%; mp 63-65 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.14-7.16 (m, 2H), 6.91 (dd, *J* = 8.2, 2.7 Hz, 1H), 4.92 (s, 1H), 4.37 (s, 2H), 2.43 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 194.1, 153.3, 135.4, 133.4, 131.5 119.4, 115.7, 33.5, 20.3; IR (film): $\tilde{\upsilon}$ = 3232, 2962, 2924, 1662, 1562, 1492, 1294, 1198, 824, 632 cm^{-1}; MS (GC/MS): *m/z*: calcd for $C_9H_{10}BrO_2$: 230.1, found: 230.2 $[M + H]^+$.

General procedure for the preparation of alkylated xanthines 26-40, 42, 44

3-Alkyl-8-bromoxanthine (0.55 mmol) was dissolved in anhydrous DMF (4 mL), and distilled *N,N*-diisopropylethylamine (1.5 eq) was added. After stirring for 5 min at room temperature, α-bromoacetophenone (1 eq) was slowly added. The reaction was stirred at room temperature for 1-3 days. The mixture was then poured in a mixture EtOAc/HCl (10%), and the formed precipitate was filtrated, washed with water to afford **26-42** in pure form.

3-Benzyl-8-bromo-3,7-dihydro-7-(2-oxo-2-phenylethyl)-1H-purine-2,6-dione (26)

White solid; Yield: 76%; mp 249-252 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.43 (s, 1H), 8.10 (d, *J* = 7.4 Hz, 2H), 7.77 (t, *J* = 7.4 Hz, 1H), 7.63 (t, *J* = 7.4 Hz, 2H), 7.27-7.38 (m, 5H), 5.94 (s, 2H), 5.11 (s, 2H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 191.2, 153.9, 150.2, 148.5, 136.5, 134.5, 133.5, 129.2, 129.0, 128.4,

128.2, 127.3, 127.3, 109.0, 52.9, 44.9; IR (film): $\tilde{\upsilon}$ = 3190, 3059, 2978, 2936, 1707, 1697, 1669, 1592, 1536, 1446, 1358, 1260, 742 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{20}H_{16}BrN_4O_3$: 439.0, found: 439.1 $[M + H]^+$.

3-Benzyl-8-bromo-7-[2-(4-fluorophenyl)-2-oxoethyl]-3,7-dihydro-1H-purine-2,6-dione (27)

White solid; Yield: 62%; mp 258-261 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.43 (s, 1H), 8.18-8.22 (m, 2H), 7.45-7.49 (m, 2H), 7.27-7.38 (m, 5H), 5.93 (s, 2H), 5.11 (s, 2H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 189.9, 153.9, 150.2, 148.6, 136.5, 131.4 (d, *J* = 9.8 Hz), 130.3 (d, *J* = 2.7 Hz), 129.2, 128.4, 127.5, 127.3, 127.3, 116.1 (d, *J* = 22.1 Hz), 109.0, 52.8, 44.9; IR (film): $\tilde{\upsilon}$ = 3184, 3068, 2972, 2934, 1697, 1665, 1592, 1536, 1359, 1234, 834, 742 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{20}H_{14}BrFN_4NaO_3$: 479.0131, found: 479.0137 $[M + Na]^+$.

8-Bromo-3,7-dihydro-3-methyl-7-[2-(4-nitrophenyl)-2-oxoethyl]-1H-purine-2,6-dione (28)

Yellow solid; Yield: 63%; mp 290-292 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.32 (s, 1H), 8.42 (d, *J* = 8.9 Hz, 2H), 8.34 (d, *J* = 8.9 Hz, 2H), 6.01 (s, 2H), 3.37 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 190.9, 153.9, 150.6, 150.4, 149.0, 138.2, 129.8, 129.0, 124.0, 108.8, 53.2, 28.5; IR (film): $\tilde{\upsilon}$ = 3158, 3037, 2939, 2831, 1674, 1523, 1369, 1346, 1220, 1203, 852 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{14}H_{11}BrN_5O_5$: 407.9, found: 408.1 $[M + H]^+$.

7-[2-(Benzo[*d*][1,3]-dioxol-5-yl)-2-oxoethyl]-8-bromo-3,7-dihydro-3-methyl-1H-purine-2,6(3H,7H)-dione (29)

White solid; Yield: 61%; mp 297-299 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.23 (s, 1H), 7.76 (dd, *J* = 8.2, 1.5 Hz, 1H), 7.56 (d, *J* = 1.5 Hz, 1H), 7.13 (d, *J* = 8.2 Hz, 1H), 6.19 (s, 2H), 5.83 (s, 2H), 3.37 (s, 3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 189.0, 153.9, 152.5, 150.4, 148.9, 148.1, 129.1, 127.9, 125.0, 108.9, 108.4, 107.3, 102.3, 52.6, 28.5; IR (film): $\tilde{\upsilon}$ = 3155, 3039, 2942, 2811, 1674, 1447, 1357, 1258, 1035, 930, 869, 817 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{15}H_{11}BrN_4NaO_5$: 428.9811, found: 428.9808 $[M + Na]^+$.

7-[2-(Benzo[*d*][1,3]-dioxol-4-yl)-2-oxoethyl]-8-bromo-3,7-dihydro-3-methyl-1H-purine-2,6-dione (30)

White solid; Yield: 67%; mp 306-308 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.31 (s, 1H), 7.35 (d, *J* = 7.9 Hz, 1H), 7.27 (d, *J* = 7.9 Hz, 1H), 7.02 (t, *J* = 7.9 Hz, 1H), 6.28 (s, 2H), 5.69 (s, 2H), 3.36 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 188.1, 153.9, 150.4, 148.9, 148.7, 148.4, 129.0, 122.0, 119.9, 115.9, 113.6, 108.9, 102.4, 55.0, 28.5; IR (film): $\tilde{\upsilon}$ = 3170, 3057, 2831, 1680, 1454, 1352, 1235, 1064, 946,745 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{15}H_{11}BrN_4NaO_5$: 428.9811, found: 428.9808 $[M + Na]^+$.

8-Bromo-3,7-dihydro-7-[2-(2-methoxyphenyl)-2-oxoethyl]-3-methyl-1H-purine-2,6-dione (31)

White solid; Yield: 75%; mp 294-296 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.29 (s, 1H), 7.78 (d, *J* = 7.4 Hz, 1H), 7.69 (t, *J* = 7.4 Hz, 1H), 7.30 (d, *J* = 7.4 Hz, 1H), 7.12 (t, *J* = 7.4 Hz, 1H), 5.68 (s, 2H), 4.02 (s, 3H),

3.36 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 191.1, 159.6, 153.9, 150.4, 148.9, 135.9, 130.2, 129.0, 123.2, 120.8, 112.8, 108.8, 56.4, 56.1, 28.5; IR (film): $\tilde{\upsilon}$ = 3170, 2987, 2900, 1673, 1462, 1357, 1204, 758 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{15}H_{13}BrN_4NaO_4$: 415.0018, found: 415.0019 $[M + Na]^+$.

8-Bromo-3,7-dihydro-7-[2-(3-methoxyphenyl)-2-oxoethyl]-3-methyl-1H-purine-2,6-dione (32)

White solid; Yield: 75%; mp 263-265 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 11.30 (s, 1H), 7.71 (ddd, *J* = 7.7, 1.5, 0.9 Hz, 1H), 7.57 (dd, *J* = 2.6, 1.5 Hz, 1H), 7.54 (dd, *J* = 8.1, 7.7 Hz, 1H), 7.33 (ddd, *J* = 8.1, 2.6, 0.9 Hz, 1H), 5.92 (s, 2H), 3.86 (s, 3H), 3.37 (s, 3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 191.0, 159.5, 153.9, 150.4, 148.9, 134.8, 130.2, 129.1, 120.7, 120.6, 112.5, 108.9, 55.4, 53.0, 28.5; IR (film): $\tilde{\upsilon}$ = 3145, 3027, 2807, 1679, 1535, 1365, 1261, 748 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{15}H_{13}BrN_4NaO_4$: 415.0018, found: 415.0018 $[M + Na]^+$.

8-Bromo-3,7-dihydro-7-[2-(4-methoxyphenyl)-2-oxoethyl]-3-methyl-1H-purine-2,6-dione (33)

White solid; Yield: 73%; mp 280-282 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.29 (s, 1H), 8.08 (d, *J* = 8.9 Hz, 2H), 7.13 (d, *J* = 8.9 Hz, 2H), 5.86 (s, 2H), 3.88 (s, 3H), 3.37 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 189.2, 164.1, 153.9, 150.4, 148.9, 130.6, 129.1, 126.4, 114.3, 108.9, 55.6, 52.5, 28.5; IR (film): $\tilde{\upsilon}$ = 3153, 3034, 2941, 2838, 1680, 1600, 1364, 1173, 834 cm^{-1}; MS (ESI): *m/z*: calcd for $C_{15}H_{14}BrN_4O_4$: 393.2, found: 393.1 $[M + H]^+$.

8-Bromo-3,7-dihydro-3-methyl-7-[2-(2-methylphenyl)-2-oxoethyl]-1H-purine-2,6-dione (34)

White solid; Yield: 58%; mp 242-244 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.32 (s, 1H), 8.04 (dd, *J* = 7.6, 1.1 Hz, 1H), 7.56 (dt, *J* = 7.6, 1.1 Hz, 1H), 7.44 (t, *J* = 7.6 Hz, 1H), 7.40 (d, *J* = 7.6 Hz, 1H), 5.79 (s, 2H), 3.37 (s, 3H), 2.42 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 194.1, 153.9, 150.4, 148.9, 138.3, 133.7, 132.7, 131.9, 129.1, 129.0, 126.1, 108.8, 54.3, 28.5, 20.6; IR (film): $\tilde{\upsilon}$ = 3154, 3025, 2814, 1676, 1535, 1364, 1208, 768 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{15}H_{13}BrN_4NaO_3$: 399.0069, found: 399.0068 $[M + Na]^+$.

8-Bromo-3,7-dihydro-3-methyl-7-[2-(3-methylphenyl)-2-oxoethyl]-1H-purine-2,6-dione (35)

White solid; Yield: 75%; mp 280-282 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.30 (s, 1H), 7.93 (s, 1H), 7.90 (d, *J* = 7.6 Hz, 1H), 7.58 (d, *J* = 7.6 Hz, 1H), 7.51 (t, *J* = 7.6 Hz, 1H), 5.91 (s, 2H), 3.37 (s, 3H), 2.42 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 191.2, 153.9, 150.4, 148.9, 138.6, 135.1, 133.5, 129.1, 128.9, 128.5, 125.3, 108.9, 52.9, 28.5, 20.7; IR (film): $\tilde{\upsilon}$ = 3149, 3024, 2813, 1678, 1536, 1360, 1164, 715 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{15}H_{13}BrN_4NaO_3$: 399.0069, found: 399.0070 $[M + Na]^+$.

8-Bromo-3,7-dihydro-3-methyl-7-[2-(4-methylphenyl)-2-oxoethyl]-1H-purine-2,6-dione (36)

White solid; Yield: 68%; mp 293-295 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.29 (s, 1H), 8.00 (d, *J* = 7.8 Hz, 2H), 7.43 (d, *J* = 7.8 Hz, 2H), 5.89 (s, 2H), 3.37 (s, 3H), 2.42 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ =

190.6, 153.9, 150.4, 148.9, 145.2, 131.0, 129.5, 129.1, 128.2, 108.9, 52.7, 28.5, 21.2; IR (film): $\tilde{\upsilon}$ = 3155, 3002, 2819, 1677, 1538, 1369, 1203, 751 cm^{-1}; MS (ESI): *m/z*: calcd for $C_{15}H_{14}BrN_4O_3$: 377.0, found: 377.1 $[M + H]^+$.

8-Bromo-3,7-dihydro-7-[2-(2-hydroxyphenyl)-2-oxoethyl]-3-methyl-1H-purine-2,6-dione (37)

White solid; Yield: 48%; mp 289-291 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 11.28 (s, 1H), 11.19 (s, 1H), 7.80 (dd, *J* = 7.9, 1.8 Hz, 1H), 7.55 (ddd, *J* = 8.3, 7.2, 1.8 Hz, 1H), 7.08 (dd, *J* = 8.3, 0.9 Hz, 1H), 6.98 (ddd, *J* = 7.9, 7.2, 0.9 Hz, 1H), 5.76 (s, 2H), 3.37 (s, 3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 191.8, 159.2, 153.9, 150.4, 148.9, 136.1, 130.0, 129.1, 120.5, 119.5, 117.5, 108.9, 55.8, 28.5; IR (film): $\tilde{\upsilon}$ = 3150, 3034, 2807, 1680, 1535, 1366, 1208, 748 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{14}H_{11}BrN_4NaO_4$: 400.9861, found: 400.9856 $[M + Na]^+$.

8-Bromo-3,7-dihydro-7-[2-(3-hydroxyphenyl)-2-oxoethyl]-3-methyl-1H-purine-2,6-dione (38)

White solid; Yield: 75%; mp 299-305 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 11.30 (s, 1H), 9.97 (s, 1H), 7.57 (d, *J* = 8.1 Hz, 1H), 7.39-7.44 (m, 2H), 7.15 (ddd, *J* = 8.1, 2.5, 0.7 Hz, 1H), 5.86 (s, 2H), 3.37 (s, 3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 191.0, 157.8, 153.9, 150.4, 148.9, 134.8, 130.2, 129.1, 121.6, 119.1, 114.1, 108.9, 52.9, 28.5; IR (film): $\tilde{\upsilon}$ = 3370, 3008, 2822, 1676, 1534, 1364, 1209, 750 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{14}H_{11}BrN_4NaO_4$: 400.9861, found: 400.9860 $[M + Na]^+$.

8-Bromo-3,7-dihydro-7-[2-(4-hydroxyphenyl)-2-oxoethyl]-3-methyl-1H-purine-2,6-dione (39)

Yellow solid; Yield: 74%; mp 320-322 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.28 (s, 1H), 10.64 (s, 1H), 7.97 (d, *J* = 8.6 Hz, 2H), 6.93 (d, *J* = 8.6 Hz, 2H), 5.81 (s, 2H), 3.36 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 188.8, 163.2, 153.9, 150.4, 148.9, 130.8, 129.2, 125.0, 115.6, 108.9, 52.4, 28.5; IR (film): $\tilde{\upsilon}$ = 3154, 2818, 1680, 1577, 1360, 1208, 1166, 759 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{14}H_{11}BrN_4NaO_4$: 400.9861, found: 400.9858 $[M + Na]^+$.

8-Bromo-3,7-dihydro-3-methyl-7-[2-(2-methyl-3-hydroxyphenyl)-2-oxoethyl]-1H-purine-2,6-dione (40)

White solid; Yield: 72%; mp 277-279 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 11.32 (s, 1H), 9.80 (s, 1H), 7.37 (dd, *J* = 7.8, 0.8 Hz, 1H), 7.22 (t, *J* = 7.8 Hz, 1H), 7.07 (dd, *J* = 7.8, 0.8 Hz, 1H), 5.70 (s, 2H), 3.36 (s, 3H), 2.18 (s, 3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 195.1, 156.1, 153.9, 150.4, 148.9, 136.0, 129.0, 126.4, 123.6, 118.9, 118.6, 108.8, 54.7, 28.5, 12.2; IR (film): $\tilde{\upsilon}$ = 3411, 3145, 3023, 2823, 1663, 1535, 1366, 1274, 1205, 785 cm^{-1}; HRMS (ESI): *m/z*: calcd $C_{15}H_{13}BrN_4NaO_4$: 415.0018, found: 415.0011 $[M + Na]^+$.

8-Bromo-3,7-dihydro-3-methyl-7-[2-(2-methyl-5-hydroxyphenyl)-2-oxoethyl]-1H-purine-2,6-dione (42)

White solid; Yield: 35%; mp 271-274 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 11.32 (s, 1H), 9.73 (s, 1H), 7.34 (d, *J* = 2.6 Hz, 1H), 7.17 (d, *J* = 8.3 Hz, 1H), 6.96 (dd, *J* = 2.6, 8.3 Hz, 1H), 5.71 (s, 2H), 3.37 (s, 3H), 2.29 (s,

3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 194.1, 155.3, 153.9, 150.4, 148.9, 134.5, 132.9, 129.1, 127.9, 119.6, 115.4, 108.8, 54.3, 28.5, 19.6; IR (film): $\tilde{\upsilon}$ = 3370, 3145, 3025, 2829, 1677, 1535, 1365, 1293, 1183, 749 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{15}H_{13}BrN_4NaO_4$: 415.0018, found: 415.0007 $[M + Na]^+$.

8-Bromo-7-[2-(4-cyanophenyl)-2-oxoethyl]-3,7-dihydro-3-methyl-1H-purine-2,6-dione (44)

White solid; Yield: 79%; mp 281-284 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.31 (s, 1H), 8.25 (d, *J* = 8.5 Hz, 2H), 8.11 (d, *J* = 8.5 Hz, 2H), 5.98 (s, 2H), 3.36 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 191.1, 153.9, 150.4, 149.0, 136.7, 132.9, 129.0, 128.9, 117.9, 116.3, 108.8, 53.1, 28.5; IR (film): $\tilde{\upsilon}$ = 3162, 3043, 2963, 2932, 1705, 1671, 1540, 1369, 1204, 834 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{15}H_{11}BrN_5O_3$: 388.0, found: 388.0 $[M + H]^+$.

8-Bromo-3,7-dihydro-3-methyl-7-[2-(2-methyl-4-hydroxyphenyl)-2-oxoethyl]-1H-purine-2,6-dione (41)

8-Bromo-3,9-dihydro-3-methyl-1H-purine-2,6-dione (**13**, 400 mg, 1.41 mmol) was added to a solution of KOH (79 mg, 1.41 mmol) in EtOH (3 mL). The resulting mixture was then heated to reflux for 2 h. EtOH was then removed under reduced pressure, and the resulting solid was washed with cold EtOH, and filtered off to afford a light yellow solid. A mixture of this solid (250 mg, 0.88 mmol) and α-bromo-4-hydroxy-2-methylacetophenone (202 mg, 0.88 mmol) in anhydrous DMF (2.5 mL) was stirred at room temperature for 16 h. The mixture was then poured in a mixture EtOAc/HCl (10%), and the formed precipitate was filtrated off, washed with water to afford the desired product as a white solid (177 mg, 51% yield). mp 313-315 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 11.28 (s, 1H), 10.45 (s, 1H), 8.00 (d, *J* = 8.6 Hz, 1H), 6.78 (dd, *J* = 8.6, 2.4 Hz, 1H), 6.73 (d, *J* = 2.4 Hz, 1H), 5.73 (s, 2H), 3.36 (s, 3H), 2.39 (s, 3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 190.6, 161.6, 153.9, 150.4, 148.9, 142.6, 132.6, 129.2, 124.2, 118.9, 112.8, 108.9, 53.6, 28.5, 21.8; IR (film): $\tilde{\upsilon}$ = 3225, 3145, 3025, 2987, 1673, 1540, 1371, 1204, 1065, 743 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{15}H_{13}BrN_4NaO_4$: 415.0018, found: 415.0015 $[M + Na]^+$.

8-Bromo-7-[2-(3,4-dihydroxyphenyl)-2-oxoethyl]-3,7-dihydro-3-methyl-1H-purine-2,6-dione (43)

8-Bromo-3,9-dihydro-3-methyl-1H-purine-2,6-dione (**13**, 500 mg, 2.04 mmol) was dissolved in anhydrous DMF (5 mL), and distilled diisopropylethylamine (0.534 mL, 3.06 mmol) was added. After stirring at 5 min at room temperature, α-chloro-3,4-dihydroxyacetophenone (380 mg, 2.04 mmol) was slowly added. The reaction mixture was heated to 70 °C for 5 h. The mixture was then poured in a mixture EtOAc/HCl (10%), and the organic layer was concentrated under reduced pressure. The resulting suspension was filtered, washed with MeOH to afford the desired product as a white solid (294 mg, 36% yield). mp 311-314 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 11.29 (s, 1H), 10.16 (s, 1H), 9.52 (s, 1H), 7.53 (dd, *J* = 8.2, 1.9 Hz, 1H), 7.40 (d, *J* = 1.9 Hz, 1H), 6.90 (d, *J* = 8.2 Hz, 1H), 5.77 (s, 2H), 3.36 (s, 3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 188.8, 153.9, 151.9, 150.4, 148.9, 145.6, 129.2, 125.3, 121.7, 115.3, 114.7,

108.9, 52.3, 28.5; IR (film): $\tilde{\upsilon}$ = 3460, 3233, 3044, 2945, 1698, 1666, 1357, 1289, 1186, 1065, 763 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{14}H_{11}BrN_4NaO_5$: 416.9808, found: 416.9805 $[M + Na]^+$.

General procedure for the cyclization of alkylated xanthines 45-50, 54, 55, 58, 60, 62, 64-68

A mixture of 3-alkyl-8-bromo-3,7-dihydro-7-(2-oxo-2-phenyl-ethyl)-1H-purine-2,6-dione (1.0 eq) and the corresponding primary amine (4.0 eq) in EtOH (0.1 M) was heated in a sealed tube at 175 °C for 12 h. The reaction was cooled to room temperature, the solid was filtered off, and washed with water affording the corresponding products in pure form. This method was used to obtain compounds **45-50**, **54**, **55**, **58**, **60**, **62**, **64-68**.

General procedure for the cyclization of alkylated xanthines 51-53, 56, 57, 59, 61, 63

A mixture of 3-alkyl-8-bromo-3,7-dihydro-7-(2-oxo-2-phenyl-ethyl)-1H-purine-2,6-dione (1.0 eq) and butylamine (4.0 eq) in *n*-PrOH (0.1 M) was heated in a sealed tube at 175 °C for 12 h. The reaction was cooled to room temperature, the solid was filtered off, and washed with water. Subsequent recristallization in EtOH afforded the corresponding products in pure form. This method was used to obtain compounds **51-53**, **56**, **57**, **59**, **61**, **63**.

1-Benzyl-8-(2-methoxyphenyl)-7-phenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (45)

White solid; Yield: 40%; mp 301-303 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 11.08 (s, 1H), 8.00 (s, 1H), 7.46-7.50 (m, 2H), 7.23-7.29 (m, 10H), 7.18 (d, *J* = 7.8 Hz, 1H), 7.06 (t, *J* = 7.8 Hz, 1H), 5.06 (s, 2H), 3.54 (s, 3H); ^{13}C NMR (100 MHz, DMSO-d_6, $CDCl_3$): δ = 154.7, 153.3, 152.4, 150.6, 147.8, 136.9, 132.9, 131.1, 129.5, 128.3, 128.2, 128.1, 127.5, 127.1, 127.1, 122.4, 120.9, 113.0, 105.4, 99.1, 55.6, 44.9, 1C missing due to overlapping; IR (film): $\tilde{\upsilon}$ = 3151, 3030, 2817, 1669, 1488, 1263, 1155, 1024, 754, 529 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{27}H_{21}N_5NaO_3$: 486.1542, found: 486.1540 $[M + Na]^+$.

1-Benzyl-7-*p*-fluorophenyl-8-(2-methoxyphenyl)-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (46)

White solid; Yield: 19%; mp 251-253 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.09 (s, 1H), 7.99 (s, 1H), 7.46-7.49 (m, 2H), 7.25-7.29 (m, 6H), 7.20-7.23 (m, 1H), 7.12-7.17 (m, 3H), 7.04-7.08 (m, 1H), 5.05 (s, 2H), 3.54 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 161.9 (d, *J* = 246.7 Hz), 154.7, 153.4, 152.4, 150.7, 147.8, 136.9, 132.0, 131.3, 129.9 (d, *J* = 8.4 Hz), 129.6, 128.3, 127.1, 127.1, 124.6 (d, *J* = 3.1 Hz), 122.2, 121.0, 115.4 (d, *J* = 21.8 Hz), 113.1, 105.6, 99.1, 55.7, 44.9; IR (film): $\tilde{\upsilon}$ = 3160, 3025, 2822, 1671, 1487, 1225, 1157, 841, 744, 529, 409 cm^{-1}; MS (ESI): *m/z*: calcd for $C_{27}H_{20}FN_5NaO_3$: 504.4, found: 505.0 $[M + Na]^+$.

8-(2-Methoxyphenyl)-1-methyl-7-*p*-nitrophenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (47)

Yellow solid; Yield: 63%; mp 323-325 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 11.06 (s, 1H), 8.35 (s, 1H), 8.14 (d, *J* = 8.7 Hz, 2H), 7.64 (dd, *J* = 7.6, 1.1 Hz, 1H), 7.55 (dt, *J* = 7.6, 1.1 Hz, 1H), 7.51 (d, *J* = 8.7 Hz, 2H),

7.21 (d, *J* = 7.6 Hz, 1H), 7.16 (t, *J* = 7.6 Hz, 1H), 3.54 (s, 3H), 3.30 (s, 3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 154.4, 153.3, 153.2, 150.9, 148.4, 146.5, 134.8, 131.5, 130.8, 129.5, 127.7, 123.6, 122.1, 121.2, 113.1, 107.6, 99.2, 55.7, 28.8; IR (film): $\tilde{\upsilon}$ = 3012, 2826, 1680, 1501, 1016, 844, 720, 600, 432 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{21}H_{16}N_6NaO_5$: 455.1080, found: 455.1085 $[M + Na]^+$.

7-(Benzo[*d*][1,3]dioxol-5-yl)-8-(2-methoxyphenyl)-1-methyl-1H-imidazo[1,2-*f*]purine-2,4(3H,8H)-dione (48)

White solid; Yield: 33%; mp 323-328 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 10.97 (s, 1H), 7.90 (s, 1H), 7.50-7.54 (m, 2H), 7.20 (dd, *J* = 8.9, 1.2 Hz, 1H), 7.10 (ddd, *J* = 8.8, 7.7, 1.2 Hz, 1H), 6.82-6.84 (m, 2H), 6.71 (dd, *J* = 8.1, 1.8 Hz, 1H), 6.00 (s, 2H), 3.62 (s, 3H), 3.29 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 154.9, 153.3, 152.7, 150.9, 147.8, 147.4, 147.1, 132.8, 131.3, 129.9, 122.5, 121.7, 121.6, 120.9, 112.9, 108.2, 107.9, 104.9, 101.2, 99.0, 55.7, 28.8; IR (film): $\tilde{\upsilon}$ = 3142, 3007, 2820, 1674, 1484, 1448, 1235, 1024, 737 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{22}H_{17}N_5NaO_5$: 454.1127, found: 454.1125 $[M + Na]^+$.

7-(Benzo[*d*][1,3]dioxol-5-yl)-8-(4-hydroxybutyl)-1-methyl-1H-imidazo[1,2-*f*]purine-2,4(3H,8H)-dione (49)

White solid; Yield: 12%; mp 267-270 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 10.85 (s, 1H), 7.61 (s, 1H), 7.16-7.17 (m, 1H), 7.05-7.06 (m, 2H), 6.12 (s, 2H), 4.35 (s, 1H), 4.09 (t, *J* = 7.4 Hz, 2H), 3.40 (s, 3H), 3.26 (t, *J* = 7.4 Hz, 2H), 1.67 (quint, *J* = 7.4 Hz, 2H), 1.24 (quint, *J* = 7.4 Hz, 2H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 153.2, 152.7, 150.9, 147.9, 147.7, 147.6, 131.9, 123.1, 121.4, 109.3, 108.6, 104.8, 101.5, 98.8, 59.8, 43.6, 29.1, 28.8, 25.0; IR (film): $\tilde{\upsilon}$ = 3575, 3369, 3161, 3012, 2936, 2822, 1676, 1498, 1451, 1243, 1028, 742 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{19}H_{19}N_5NaO_5$: 420.1284, found: 420.1289 $[M + Na]^+$.

7-(Benzo[*d*][1,3]dioxol-4-yl)-8-(2-methoxyphenyl)-1-methyl-1H-imidazo[1,2-*f*]purine-2,4(3H,8H)-dione (50)

White solid; Yield: 38%; mp 347-349 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.01 (s, 1H), 7.87 (s, 1H), 7.51-7.53 (m, 2H), 7.20 (d, *J* = 8.5 Hz, 1H), 7.11 (t, *J* = 7.5 Hz, 1H), 6.85 (d, *J* = 7.9 Hz, 1H), 6.68 (t, *J* = 7.9 Hz, 1H), 6.44 (d, *J* = 7.9 Hz, 1H), 5.99 (s, 1H), 5.95 (s, 1H), 3.59 (s, 3H), 3.29 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 154.8, 153.3, 152.9, 150.9, 147.7, 147.2, 144.6, 131.2, 129.5, 126.9, 122.6, 121.4, 120.8, 119.9, 112.9, 110.3, 108.5, 106.7, 101.1, 99.0, 55.7, 28.8; IR (film): $\tilde{\upsilon}$ = 3520, 3151, 3000, 2822, 1667, 1600, 1507, 1438, 1025, 739 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{22}H_{17}N_5NaO_5$: 454.1127, found: 454.1113 $[M + Na]^+$.

8-(Butyl)-1-methyl-7-*o*-methoxyphenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (51)

White solid; Yield: 28%; mp 242-244 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.89 (s, 1H), 7.53 (t, *J* = 7.8 Hz, 1H), 7.50 (s, 1H), 7.38 (d, *J* = 7.8 Hz, 1H), 7.19 (d, *J* = 7.8 Hz, 1H), 7.08 (t, *J* = 7.8 Hz, 1H), 3.84 (t, *J* = 7.3

Hz, 2H), 3.81 (s, 3H), 3.39 (s, 3H), 1.60 (quint, *J* = 7.3 Hz, 2H), 1.08 (sext, *J* = 7.3 Hz, 2H), 0.69 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 157.3, 153.2, 152.7, 150.9, 147.3, 132.3, 131.5, 129.0, 120.6, 116.3, 111.4, 105.3, 98.8, 55.4, 43.4, 29.9, 28.8, 18.8, 12.9; IR (film): $\tilde{\upsilon}$ = 3145, 2999, 2934, 2837, 1674, 1514, 1461, 1246, 745 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{19}H_{21}N_5NaO_3$: 390.1542, found: 390.1545 [M + Na]$^+$.

8-(Butyl)-1-methyl-7-*m*-methoxyphenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (52)

White solid; Yield: 52%; mp 206-208 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.91 (s, 1H), 7.70 (s, 1H), 7.44 (t, *J* = 8.2 Hz, 1H), 7.13-7.14 (m, 2H), 7.05 (dd, *J* = 8.2, 2.4 Hz, 1H), 4.11 (t, *J* = 7.3 Hz, 2H), 3.82 (s, 3H), 3.39 (s, 3H), 1.63 (quint, *J* = 7.3 Hz, 2H), 1.13 (sext, *J* = 7.3 Hz, 2H), 0.74 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 159.4, 153.2, 152.7, 150.9, 147.9, 132.0, 130.0, 129.2, 120.8, 114.8, 114.1, 105.2, 98.8, 55.2, 43.5, 30.2, 28.8, 18.9, 13.1; IR (film): $\tilde{\upsilon}$ = 3141, 3044, 2957, 2817, 1679, 1509, 1467, 1156, 716 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{19}H_{21}N_5NaO_3$: 390.1542, found: 390.1544 [M + Na]$^+$.

8-(Butyl)-1-methyl-7-*p*-methoxyphenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (53)

White solid; Yield: 89%; mp 272-274 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.90 (s, 1H), 7.60 (s, 1H), 7.51 (d, *J* = 8.7 Hz, 2H), 7.09 (d, *J* = 8.7 Hz, 2H), 4.07 (t, *J* = 7.3 Hz, 2H), 3.83 (s, 3H), 3.40 (s, 3H), 1.62 (quint, *J* = 7.3 Hz, 2H), 1.12 (sext, *J* = 7.3 Hz, 2H), 0.75 (t, *J* = 7.3 Hz, 3H); IR (film): $\tilde{\upsilon}$ = 3144, 3051, 2963, 2864, 2802, 1671, 1507, 1247, 831, 742 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{19}H_{21}N_5NaO_3$: 390.1542, found: 390.1541 [M + Na]$^+$.

8-(2-Methoxyphenyl)-1-methyl-7-*o*-methylphenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (54)

White solid; Yield: 64%; mp 301-303 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 10.98 (s, 1H), 7.81 (s, 1H), 7.49 (dd, *J* = 7.7, 1.6 Hz, 1H), 7.39 (ddd, *J* = 8.6, 8.3, 1.6 Hz, 1H), 7.21-7.22 (m, 2H), 7.13-7.15 (m, 1H), 7.04-7.08 (m, 2H), 7.00 (ddd, *J* = 8.6, 7.7, 1.2 Hz, 1H), 3.57 (s, 3H), 3.31 (s, 3H), 2.27 (s, 3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 154.8, 153.3, 152.9, 150.9, 147.7, 147.2, 144.6, 131.2, 129.5, 126.9, 122.6, 121.4, 120.8, 119.9, 112.9, 110.3, 108.5, 106.7, 101.1, 99.0, 55.7, 28.8; IR (film): $\tilde{\upsilon}$ = 3166, 3007, 2811, 1666, 1490, 1157, 744 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{22}H_{19}N_5NaO_3$: 424.1386, found: 424.1389 [M + Na]$^+$.

8-(2-Methoxyphenyl)-1-methyl-7-*m*-methylphenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (55)

White solid; Yield: 72%; mp 313-315 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.98 (s, 1H), 7.95 (s, 1H), 7.51 (d, *J* = 7.2 Hz, 2H), 6.98-7.21 (m, 6H), 3.58 (s, 3H), 3.30 (s, 3H), 2.20 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 154.9, 153.3, 152.7, 150.9, 147.9, 137.5, 133.0, 131.2, 129.8, 128.9, 128.2, 128.1, 128.0, 124.3, 122.6, 120.9, 112.9, 105.2, 99.1, 55.7, 28.8, 20.8; IR (film): $\tilde{\upsilon}$ = 3160, 3028, 2822, 1704, 1671, 1505, 1435, 1160, 738 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{22}H_{19}N_5NaO_3$: 424.1386, found: 424.1383 [M + Na]$^+$.

8-(Butyl)-1-methyl-7-*m*-methylphenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (56)

White solid; Yield: 42%; mp 238-240 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.91 (s, 1H), 7.65 (s, 1H), 7.40-7.43 (m, 2H), 7.37 (d, *J* = 7.6 Hz, 1H), 7.31 (d, *J* = 7.6 Hz, 1H), 4.10 (t, *J* = 7.3 Hz, 2H), 3.40 (s, 3H), 2.38 (s, 3H), 1.62 (quint, *J* = 7.3 Hz, 2H), 1.12 (sext, *J* = 7.3 Hz, 2H), 0.74 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 153.2, 152.7, 150.9, 147.9, 138.3, 132.3, 129.6, 129.3, 128.8, 127.9, 125.8, 104.9, 98.8, 43.3, 30.1, 28.8, 20.8, 18.8, 13.0; IR (film): $\tilde{\upsilon}$ = 3160, 3044, 2949, 2802, 1672, 1509, 1147, 721 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{19}H_{21}N_5NaO_2$: 374.1593, found: 374.1584 $[M + Na]^+$.

8-(Butyl)-1-methyl-7-*p*-methylphenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (57)

White solid; Yield: 24%; mp 227-229 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.91 (s, 1H), 7.63 (s, 1H), 7.46 (d, *J* = 7.9 Hz, 2H), 7.34 (d, *J* = 7.9 Hz, 2H), 4.09 (t, *J* = 7.3 Hz, 2H), 3.40 (s, 3H), 2.38 (s, 3H), 1.62 (quint, *J* = 7.3 Hz, 2H), 1.11 (sext, *J* = 7.3 Hz, 2H), 0.74 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 153.2, 152.7, 150.9, 147.8, 138.6, 132.2, 129.5, 128.8, 125.0, 104.7, 98.8, 43.3, 30.2, 28.8, 20.7, 18.8, 13.1; IR (film): $\tilde{\upsilon}$ = 3169, 3051, 2958, 2858, 2785, 1671, 1506, 1294, 1147, 825 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{19}H_{21}N_5NaO_2$: 374.1593, found: 374.1589 $[M + Na]^+$.

8-(2-Methoxyphenyl)-1-methyl-7-*o*-hydroxyphenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (58)

Yellow solid; Yield: 50%; mp 350-352 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.95 (s, 1H), 9.75 (s, 1H), 7.68 (s, 1H), 7.40-7.44 (m, 2H), 7.10-7.14 (m, 2H), 7.00-7.04 (m, 2H), 6.81 (dd, *J* = 8.2, 0.9 Hz, 1H), 6.67 (ddd, *J* = 8.4, 7.5, 0.9 Hz, 1H), 3.58 (s, 3H), 3.30 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 155.5, 154.8, 153.3, 152.7, 150.9, 147.5, 130.7, 130.6, 130.1, 129.7, 129.4, 122.8, 120.5, 118.3, 115.5, 114.7, 112.7, 106.4, 98.9, 55.5, 28.7; IR (film): $\tilde{\upsilon}$ = 3251, 3166, 2821, 1677, 1486, 1445, 1159, 744 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{21}H_{17}N_5NaO_4$: 426.1178, found: 426.1174 $[M + Na]^+$.

8-(Butyl)-1-methyl-7-*o*-hydroxyphenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (59)

White solid; Yield: 22%; mp 209-211 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 10.88 (s, 1H), 9.99 (s, 1H), 7.48 (s, 1H), 7.35 (t, *J* = 7.8 Hz, 1H), 7.30 (d, *J* = 7.8 Hz, 1H), 7.00 (d, *J* = 7.8 Hz, 1H), 6.93 (t, *J* = 7.8 Hz, 1H), 3.92 (t, *J* = 7.3 Hz, 2H), 3.40 (s, 3H), 1.60 (quint, *J* = 7.3 Hz, 2H), 1.08 (sext, *J* = 7.3 Hz, 2H), 0.70 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 155.8, 153.2, 152.6, 150.9, 147.4, 132.2, 131.2, 129.6, 119.2, 115.7, 114.8, 105.1, 98.8, 43.4, 30.1, 28.8, 18.8, 13.0; IR (film): $\tilde{\upsilon}$ = 3138, 3044, 2953, 2870, 2811, 1683, 1509, 1200, 739 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{18}H_{19}N_5NaO_3$: 376.1386, found: 376.1383 $[M + Na]^+$.

8-(2-Methoxyphenyl)-1-methyl-7-*m*-hydroxyphenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (60)

White solid; Yield: 59%; mp 353-355 °C; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.97 (s, 1H), 9.51 (s, 1H), 7.88 (s, 1H), 7.47-7.52 (m, 2H), 7.20 (d, *J* = 8.2 Hz, 1H), 7.06-7.11 (m, 2H), 6.68-6.70 (m, 2H), 6.64 (s, 1H), 3.60 (s, 3H), 3.29 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 157.1, 154.9, 153.3, 152.7, 150.9, 147.9, 133.1,

131.1, 129.7, 129.4, 129.2, 122.6, 120.9, 118.3, 115.4, 114.4, 112.9, 105.1, 99.1, 55.7, 28.8; IR (film): $\tilde{\upsilon}$ = 3147, 3030, 2817, 1676, 1482, 1299, 1159, 758 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{21}H_{17}N_5NaO_4$: 426.1178, found: 426.1181 $[M + Na]^+$.

8-(Butyl)-1-methyl-7-*m*-hydroxyphenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (61)

White solid; Yield: 41%; mp 256-258 °C; ^1H NMR (500 MHz, DMSO-d_6): δ = 10.89 (s, 1H), 9.75 (s, 1H), 7.59 (s, 1H), 7.31 (t, *J* = 7.8 Hz, 1H), 6.96 (d, *J* = 7.8 Hz, 1H), 6.93 (s, 1H), 6.88 (d, *J* = 7.8 Hz, 1H), 4.07 (t, *J* = 7.3 Hz, 2H), 3.38 (s, 3H), 1.63 (quint, *J* = 7.3 Hz, 2H), 1.12 (sext, *J* = 7.3 Hz, 2H), 0.74 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 157.6, 153.2, 152.7, 150.9, 147.8, 132.3, 129.9, 129.0, 119.5, 116.0, 115.6, 104.7, 98.8, 43.3, 30.2, 28.8, 18.8, 13.1; IR (film): $\tilde{\upsilon}$ = 3297, 3166, 3046, 2963, 2864, 1675, 1512, 1459, 1312, 1148, 722 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{18}H_{19}N_5NaO_3$: 376.1386, found: 376.1388 $[M + Na]^+$.

8-(2-Methoxyphenyl)-1-methyl-7-*p*-hydroxyphenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (62)

White solid; Yield: 3%; mp 364-366 °C; ^1H NMR (500 MHz, DMSO-d_6): δ = 10.94 (s, 1H), 9.66 (s, 1H), 7.78 (s, 1H), 7.50 (ddd, *J* = 8.4, 7.6, 1.7 Hz, 1H), 7.47 (dd, *J* = 7.7, 1.7 Hz, 1H), 7.19 (dd, *J* = 8.4, 1.1 Hz, 1H), 7.05-7.10 (m, 3H), 6.65 (d, *J* = 8.7 Hz, 2H), 3.60 (s, 3H), 3.29 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 157.6, 155.0, 153.3, 152.5, 150.9, 147.7, 133.4, 131.1, 129.9, 129.2, 122.6, 120.9, 118.7, 115.1, 112.9, 104.1, 99.0, 55.7, 28.7; IR (film): $\tilde{\upsilon}$ = 3326, 3140, 3005, 2820, 1678, 1508, 1438, 1226, 1159, 838, 742 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{21}H_{17}N_5NaO_4$: 426.1178, found: 426.1177 $[M + Na]^+$.

8-(Butyl)-1-methyl-7-*p*-hydroxyphenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (63)

White solid; Yield: 10%; mp 277-281 °C; ^1H NMR (400 MHz, DMSO-d_6): δ = 10.89 (s, 1H), 9.85 (s, 1H), 7.52 (s, 1H), 7.36 (d, *J* = 8.7 Hz, 2H), 6.89 (d, *J* = 8.7 Hz, 2H), 4.04 (t, *J* = 7.3 Hz, 2H), 3.39 (s, 3H), 1.62 (quint, *J* = 7.3 Hz, 2H), 1.12 (sext, *J* = 7.3 Hz, 2H), 0.74 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 158.2, 153.2, 152.6, 150.9, 147.6, 132.5, 130.5, 118.3, 115.6, 104.1, 98.8, 43.1, 30.2, 28.8, 18.8, 13.1; IR (film): $\tilde{\upsilon}$ = 3283, 3160, 3043, 2973, 1676, 1510, 1226, 1148, 843, 716 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{18}H_{19}N_5NaO_3$: 376.1386, found: 376.1381 $[M + Na]^+$.

8-(2-Methoxyphenyl)-1-methyl-7-(2´-methyl-3´-hydroxyphenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (64)

Light yellow solid; Yield: 56%; mp 372-374 °C; ^1H NMR (400 MHz, DMSO-d_6): δ = 10.96 (s, 1H), 9.43 (s, 1H), 7.71 (s, 1H), 7.38-7.42 (m, 2H), 7.09 (d, *J* = 8.1 Hz, 1H), 6.99 (t, *J* = 7.6 Hz, 1H), 6.89 (t, *J* = 7.6 Hz, 1H), 6.75 (d, *J* = 8.1 Hz, 1H), 6.64 (d, *J* = 7.6 Hz, 1H), 3.60 (s, 3H), 3.30 (s, 3H), 2.01 (s, 3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 155.1, 154.7, 153.3, 152.7, 150.9, 147.4, 132.1, 130.8, 129.6, 128.3, 125.4, 124.5, 122.2, 121.9, 120.5, 115.3, 112.5, 106.1, 99.0, 55.5, 28.7, 12.9; IR (film): $\tilde{\upsilon}$ = 3326, 3145, 3062, 2963, 2817,

1669, 1508, 1280, 1157, 747 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{22}H_{19}N_5NaO_4$: 440.1335, found: 440.1332 $[M + Na]^+$.

8-(2-Methoxyphenyl)-1-methyl-7-(2´-methyl-4´-hydroxyphenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (65)

White solid; Yield: 52%; mp 330-333 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 10.94 (s, 1H), 9.55 (s, 1H), 7.66 (s, 1H), 7.38-7.43 (m, 2H), 7.09 (d, *J* = 7.5 Hz, 1H), 7.00 (t, *J* = 7.5 Hz, 1H), 6.94 (d, *J* = 8.3 Hz, 1H), 6.57 (s, 1H), 6.44 (d, *J* = 8.3 Hz, 1H), 3.60 (s, 3H), 3.30 (s, 3H), 2.14 (s, 3H); IR (film): $\tilde{\upsilon}$ = 3208, 3048, 2942, 2838, 1675, 1596, 1455, 1205, 1153, 747 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{22}H_{19}N_5NaO_4$: 440.1335, found: 440.1330 $[M + Na]^+$.

8-(2-Methoxyphenyl)-1-methyl-7-(2´-methyl-5´-hydroxyphenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (66)

White solid; Yield: 45%; mp 350-352 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 10.96 (s, 1H), 9.22 (s, 1H), 7.74 (s, 1H), 7.39-7.44 (m, 2H), 7.10 (dd, *J* = 8.4, 0.9 Hz, 1H), 6.97-7.04 (m, 2H), 6.62 (dd, *J* = 8.3, 2.6 Hz, 1H), 6.56 (d, *J* = 2.6 Hz, 1H), 3.61 (s, 3H), 3.30 (s, 3H), 2.12 (s, 3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 154.6, 154.4, 153.3, 152.7, 150.9, 147.4, 131.8, 130.8, 130.7, 129.5, 127.8, 127.6, 122.2, 120.5, 117.7, 116.2, 112.6, 106.1, 99.0, 55.5, 28.7, 18.6; IR (film): $\tilde{\upsilon}$ = 3296, 3145, 1698, 1670, 1593, 1505, 1282, 1153, 1017, 751 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{22}H_{19}N_5NaO_4$: 440.1335, found: 440.1333 $[M + Na]^+$.

8-(2-Methoxyphenyl)-1-methyl-7-(2´,5´-dihydroxyphenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (67)

White solid; Yield: 34%; mp 333-336 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 10.94 (s, 1H), 9.18 (s, 1H), 8.96 (s, 1H), 7.69 (s, 1H), 7.50 (t, *J* = 7.6 Hz, 1H), 7.42 (d, *J* = 7.6 Hz, 1H), 7.20 (d, *J* = 7.6 Hz, 1H), 7.07 (t, *J* = 7.6 Hz, 1H), 6.61-6.63 (m, 2H), 6.53 (dd, *J* = 7.8, 1.8 Hz, 1H), 3.64 (s, 3H), 3.28 (s, 3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 155.1, 153.3, 152.5, 150.9, 147.7, 145.9, 144.9, 133.7, 131.1, 129.9, 122.7, 120.8, 119.4, 118.9, 115.4, 115.3, 112.9, 103.9, 99.0, 55.7, 28.8; IR (film): $\tilde{\upsilon}$ = 3446, 3243, 3062, 1714, 1656, 1506, 1464, 1276, 1148, 749 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{21}H_{17}N_5NaO_5$: 442.1121, found: 442.1122 $[M + Na]^+$.

8-(2-Methoxyphenyl)-1-methyl-7-*p*-cyanophenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (68)

White solid; Yield: 68%; mp 303-305 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 11.04 (s, 1H), 8.28 (s, 1H), 7.77 (d, *J* = 8.4 Hz, 2H), 7.62 (dd, *J* = 7.6, 1.2 Hz, 1H), 7.55 (dt, *J* = 7.6, 1.2 Hz, 1H), 7.43 (d, *J* = 8.4 Hz, 2H), 7.20 (dd, *J* = 7.6, 1.2 Hz, 1H), 7.15 (dt, *J* = 7.6, 1.2 Hz, 1H), 3.53 (s, 3H), 3.30 (s, 3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 154.4, 153.3, 153.1, 150.9, 148.3, 132.9, 132.3, 131.5, 131.1, 129.5, 127.4, 122.2, 121.2, 118.3,

113.1, 110.5, 107.2, 99.2, 55.6, 28.8; IR (film): $\tilde{\upsilon}$ = 3010, 2835, 1680, 1497, 745, 605, 445 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{22}H_{16}N_6NaO_3$: 435.1182, found: 435.1187 $[M + Na]^+$.

7-*p*-Fluorophenyl-8-(2-methoxyphenyl)-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (69)

A mixture of **46** (528 mg, 1.09 mmol), 10% Pd-C (339 mg) and dry ammonium formate (692 mg, 10.9 mmol) in absolute methanol (16 mL) was heated in a sealed tube at 140°C for 3 h. The reaction was cooled to room temperature and the mixture was filtered over a short path of celite. The residue was evaporated under reduced pressure affording the desired compound as a white solid (121 mg, 28% yield). mp 330-333 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 11.53 (s, 1H), 10.71 (s, 1H), 7.96 (s, 1H), 7.47-7.51 (m, 2H), 7.28-7.31 (m, 2H), 7.13-7.18 (m, 3H), 7.06-7.10 (m, 1H), 3.56 (s, 3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 162.4 (d, *J* = 246.1 Hz), 155.2, 154.7, 152.7, 151.7, 148.9, 132.3, 131.7, 130.2 (d, *J* = 8.4 Hz), 130.1, 125.3 (d, *J* = 3.2 Hz), 122.9, 121.4, 115.9 (d, *J* = 21.8 Hz), 113.4, 105.9, 99.5, 56.2; IR (film): $\tilde{\upsilon}$ = 3159, 2900, 1668, 1500, 1393, 1242, 1065, 750, 536 cm^{-1}; MS (ESI): *m/z*: calcd for $C_{20}H_{14}FN_5NaO_3$: 414.3, found: 414.2 $[M + Na]^+$.

8-(2-Methoxyphenyl)-1-methyl-7-*p*-carboxyphenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (70)

To a solution of 8-(2-methoxyphenyl)-1-methyl-7-*p*-cyanophenyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (**68**, 85 mg, 0.206 mmol) in water (0.910 mL) was added conc. H_2SO_4 (0.803 mL) and the mixture was refluxed for 2 h. The reaction was cooled to room temperature, water (1 mL) was added, and the precipitate formed was filtered off, affording the desired product as a white solid (60 mg, 67% yield). ^{1}H NMR (400 MHz, DMSO-d_6): δ = 11.02 (s, 1H), 8.16 (s, 1H), 7.82 (d, *J* = 8.3 Hz, 2H), 7.58 (dd, *J* = 7.7, 1.4 Hz, 1H), 7.53 (dt, *J* = 7.7, 1.4 Hz, 1H), 7.37 (d, *J* = 8.3 Hz, 2H), 7.19 (d, *J* = 7.6 Hz, 1H), 7.14 (t, *J* = 7.6 Hz, 1H), 3.54 (s, 3H), 3.30 (s, 3H), CO_2H not observed; ^{13}C NMR (100 MHz, DMSO-d_6): δ = 166.6, 154.6, 153.3, 152.9, 150.9, 148.2, 132.4, 131.9, 131.3, 130.3, 129.6, 129.2, 127.0, 122.4, 121.1, 113.0, 106.4, 99.1, 55.6, 28.8; IR (film): $\tilde{\upsilon}$ = 3013, 2839, 1681, 1547, 1264, 1158, 1013, 745, 570 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{22}H_{17}N_5NaO_5$: 454.1127, found: 454.1135 $[M + Na]^+$. HPLC purity 88%.

2.5.5 FRET based enzymatic assay

Compounds were tested in the Z'-LYTE™ Kinase Assay Kit–Tyr 1 Peptide (Invitrogen, USA) in a Corning 384 well microtiter plate. Fluorescence progress curves were measured upon excitation at 400 nm and emission at 445 and 520 nm. The assay contained a final concentration of EphB4 and ATP of 25 ng/µL and 125 µM (which is near its K_m), respectively, and was run at 30 °C for 2 h. IC_{50} values (inhibitor concentration at which enzyme activity is reduced by 50%) are determined after carrying out assays at ten different concentrations between 20 µM and 10 pM.

2.5.6 [γ-^{33}P]ATP based enzymatic assay

The enzymatic assays for the selectivity profile were performed at Reaction Biology Corporation and at the National Centre for Protein Kinase Profiling at the University of Dundee. All assays (25.5 µL volume) performed at the University of Dundee were carried out robotically at room temperature and were linear with respect to time and enzyme concentration under the conditions used. Assays were performed for 30 min using Multidrop Micro reagent dispensers (Thermo Electron Corporation, Waltham, MA, USA) in a 96-well format. The concentration of magnesium acetate in the assays was 10 mM and [γ-^{33}P]ATP (800 c.p.m./pmol) was used at 5, 20 or 50 µM in order to be at or below the K_m for ATP for each kinase. The assays were initiated with MgATP, stopped by the addition of 5 µL of 0.5 M orthophosphoric acid and spotted on to P81 filter plates using a unifilter harvester (PerkinElmer, Boston, MA, USA). IC_{50} values were determined after carrying out assays at ten different concentrations between 10 µM and 110 pM. The data is presented as mean percentage activity of duplicate assays at single concentration compared to DMSO controls. A similar protocol was used at Reaction Biology Corporation to measure the inhibitory activity of compound **66** against 11 Eph kinases.

2.5.7 Phage display based binding assay

The experiments were performed at Ambit Biosciences Corporation using binding assays as previously described.[11] Briefly, kinases were expressed as fusion proteins to T7 phage. In general, full-length constructs were used for small kinases and catalytic domains for large kinases. T7-kinase-tagged phage strains were mixed with known kinase inhibitors immobilized on streptavidin-coated magnetic beads and with test compounds at a single concentration of 1 µM. Test compounds that bind to the kinase ATP site displace the immobilized ligand from the kinase/phage, which is detected using quantitative PCR. The results are reported as the percentage of kinase/phage remaining bound to the ligand/beads, relative to a control (DMSO lacking a test compound). High affinity compounds have %control = 0, while weaker binders have higher %control values. Results are reported for screening against 50 human kinases.

2.6 Additional experimental data

Table 4. Local selectivity of compound **66**. These IC_{50} values were measured at the University of Dundee.

Kinases	Percentage of kinase activity compared to a 100% DMSO control	
	54[a]	**66**[b]
MKK1	70	70
ERK1	97	96
ERK2	91	93
JNK1	106	96
JNK2	93	94
p38a MAPK	87	97
p38b MAPK	93	83
p38g MAPK	92	84
p38d MAPK	94	86
ERK8	109	97
RSK1	83	88
RSK2	105	104
PDK1	109	103
PKBa	97	98
PKBb	104	104
SGK1	124	78
S6K1	98	77
PKA	92	96
ROCK 2	94	77
PRK2	95	93
PKCa	97	92
PKCz	98	95
PKD1	100	92
MSK1	95	95
MNK1	93	94
MNK2	85	74
MAPKAP-K2	96	90
PRAK	87	81

Kinases	Percentage of kinase activity compared to a 100% DMSO control	
	54[a]	**66**[b]
CAMKKb	101	104
CAMK1	93	100
SmMLCK	91	87
PHK	92	95
CHK1	93	95
CHK2	99	90
GSK3b	102	89
CDK2-Cyclin A	97	90
PLK1	89	93
Aurora B	104	93
AMPK	92	86
MARK3	111	96
BRSK2	81	82
MELK	89	80
CK1	108	93
CK2	85	80
DYRK1A	99	99
DYRK2	94	86
DYRK3	103	96
NEK2a	102	123
NEK6	103	89
IKKb	92	86
PIM1	96	88
PIM2	115	93
PIM3	111	81
SRPK1	78	94
MST2	95	111
EF2K	93	96
HIPK2	114	98

Kinases	Percentage of kinase activity compared to a 100% DMSO control	
	54[a]	**66**[b]
PAK4	93	102
PAK5	94	94
PAK6	90	92
Src	12	9
Lck	3	0
CSK	72	31
FGF-R1	90	62
IRR	110	100
EPHA2	10	7
MST4	83	78
SYK	94	94
YES1	6	3
IGF-1R	106	94
VEG-FR	106	77
BTK	52	33
IR-HIS	103	104
EPHB3	28	3
TBK1	79	89
IKKe	101	85
GCK	109	22
IRAK4	93	82
NUAK1	140	133
MLK1	85	45
MINK1	101	102
MLK3	136	101
LKB1	89	75
HER4	45	32
TTK	99	98

[a] Percentage of activity @ 3μM concentration of compound **54**
[b] Percentage of activity @ 1μM concentration of compound **66**

Table 5. Selectivity of compound **66** on kinases predicted to have a small gatekeeper. These IC_{50} values were measured at Ambit Biosciences Corporation.

Kinases	Percentage of kinase remaining bound to the beads, relative to a 100% DMSO control[a]
ABL1-phosphorylated	0.15
ABL2	0.75
ALK	99
BLK	0.6
BMPR1A	51
BMPR1B	0.9
BMX	55
BRK	12
DDR1	1.4
DDR2	21
EGFR	8.6
EPHA6	40
EPHB4	1.2
EPHB6	32
ERBB2	27
ERBB3	78
FGFR2	94
FGFR3	89
FGFR4	100
FLT4	93
FRK	0.3
FYN	0.25
GAK	49
GCN2(Kin.Dom.2,S808G)	100
HCK	0.6
KIT	88
LIMK1	83

Kinases	Percentage of kinase remaining bound to the beads, relative to a 100% DMSO control[a]
LIMK2	78
LYN	0.4
NEK11	67
NLK	83
PDGFRA	84
PDGFRB	53
QSK	32
RAF1	14
RET	79
RIPK2	2.9
RSK3(Kin.Dom.2-C-terminal)	84
RSK4(Kin.Dom.2-C-terminal)	5.8
SIK	41
SRMS	18
TEC	100
TESK1	9.8
TGFBR1	3.6
TGFBR2	2.4
VEGFR2	96
YANK1	80
YANK2	89
YANK3	100
ZAK	15

[a]Percentage of activity @ 1 μM concentration of compound 66

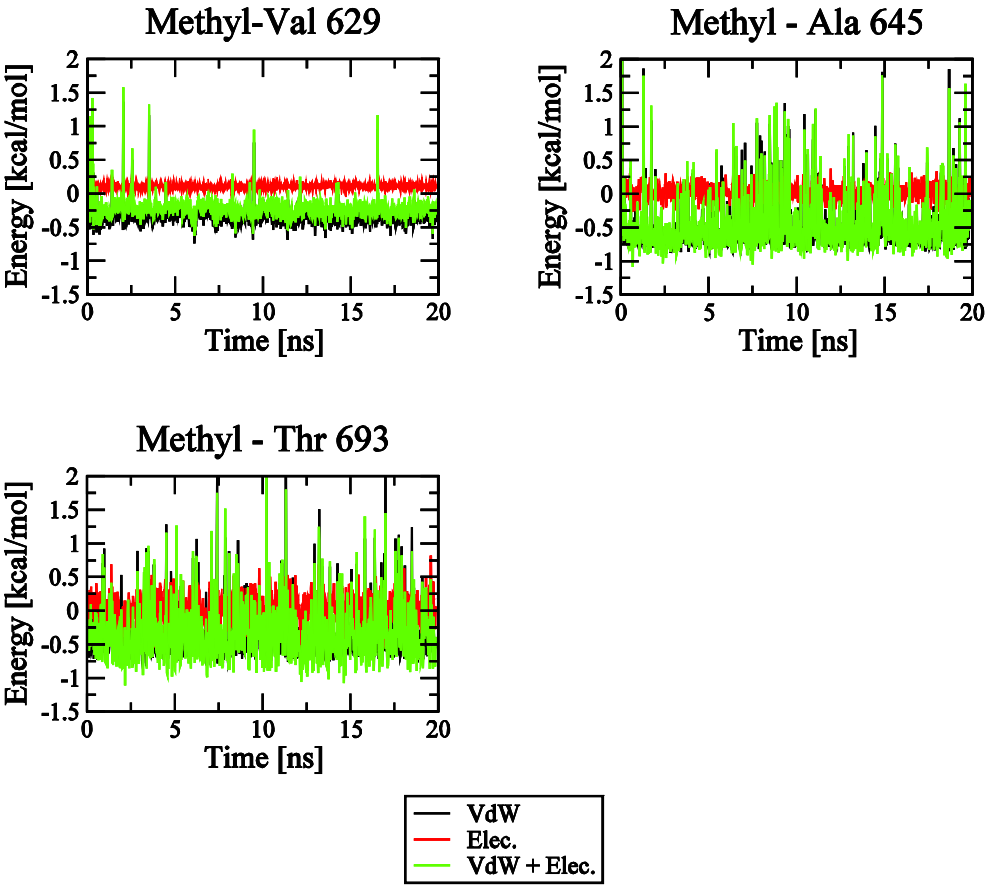


Figure 7. Time series of the van der Waals and electrostatic interaction energies between the methyl of compound **66** and the side chains of Val629, Ala645 and Thr693.

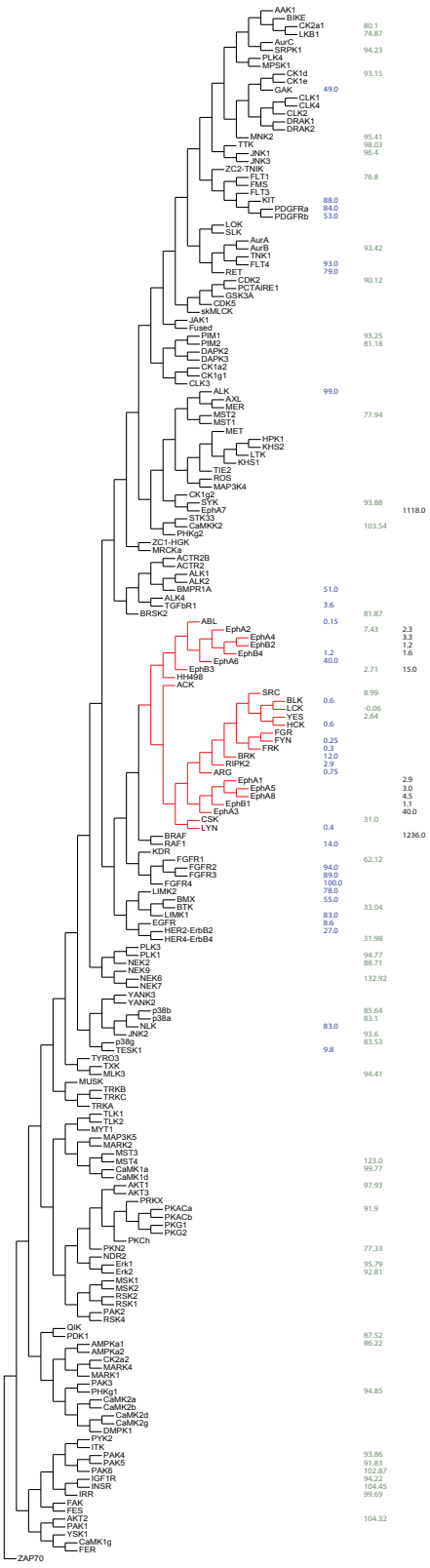


Figure 8. SAR-based dendrogram. Pourcentage of kinase activity compared to a 100% DMSO control. These values were measured at Reaction Biology Corporation (black), at the University of Dundee (green) and at Ambit Biosciences Corporation (blue).

6-Amino-1-methyl-2,4(1H,3H)-pyrimidinedione (5)

A mixture of methylurea (22.4 g, 302 mmol), 2-cyanoacetic acid (38.0 g, 447 mmol) and Ac_2O (74 mL) was heated to 60 °C for 2.5 h. After cooling to room temperature, water (74 mL) was added, and the suspension was stirred for 1 h. The solid intermediate was filtered, suspended in NaOH (20%, 150 mL), and heated to 90 °C for 1 h. The reaction mixture was then cooled to 20 °C, acidified with concentrated HCl, and the precipitate was collected by filtration (27.2 g, 64% yield) as a light yellow solid. ^{1}H NMR (400 MHz, DMSO-d_6): δ = 10.43 (s, 1H), 6.88 (s, 2H), 4.61 (s, 1H), 3.18 (s, 3H); IR (film): $\tilde{\upsilon}$ = 3344, 3183, 3101, 3046, 1660, 1618, 1548, 1276, 1161, 827 cm^{-1}; MS (ESI): *m/z*: calcd for $C_5H_7N_3NaO_2$: 164.0, found: 164.0 $[M + Na]^+$.

6-Amino-1-(phenylmethyl)-2,4(1H,3H)-pyrimidinedione (6)

A mixture of benzylurea (5.0 g, 33.3 mmol), 2-cyanoacetic acid (2.6 g, 36.63 mmol) and Ac_2O (20 mL) was heated at 80 °C for 2 h. After cooling to room temperature, Et_2O (30 mL) was added, and the resultant suspension was stirred for 1 h. The solid intermediate was filtered, suspended in a mixture of water (30 mL) and ethanol (15 mL) and heated at 85 °C. NaOH (10%, 5 mL) were slowly added. After 1 h, the reaction mixture was concentrated, acidified with concentrated HCl, and the precipitate collected by filtration. If necessary, the solid was washed with acetone to afford a light yellow solid (3.5 g, 49% yield). ^{1}H NMR (400 MHz, DMSO-d_6): δ = 10.44 (s, 1H), 7.19-7.36 (m, 5H), 6.75 (s, 2H), 5.02 (s, 2H), 4.59 (d, *J* = 2.0 Hz, 1H); ^{13}C NMR (75 MHz, DMSO-d_6): δ = 162.3, 155.7, 151.4, 136.6, 128.4, 127.1, 126.2, 75.5, 43.5; IR (film): $\tilde{\upsilon}$ = 3470, 3327, 3250, 3217, 3070, 2968, 2766, 1694, 1630, 1575, 1494, 1384, 1278, 895, 821, 728 cm^{-1}; MS (ESI): *m/z*: calcd for $C_{11}H_{11}N_3NaO_2$: 240.1, found: 240.1 $[M + Na]^+$.

General procedure for the preparation of 1-alkyl-5-nitroso-4-aminouracil (7, 8)

To a stirred mixture containing 1-alkyl-6-aminouracil (1 eq) in water (0.45 M) and acetic acid (8.2 M) at -5 ºC, $NaNO_2$ (1.1 eq) was added in small portions. A deep purple color, characteristic of the nitroso derivative appeared immediately, and stirring was continued at room temperature for 19 h. The solid was filtered off, washed with cold water, and dried under high vacuum to obtain the corresponding products in pure form.

6-Amino-1-methyl-5-nitroso-2,4(1H,3H)-pyrimidinedione (7)

Yield: 77%; ^{1}H NMR (300 MHz, DMSO-d_6): δ = 13.02 (s, 1H), 11.51 (s, 1H), 9.14 (s, 1H), 3.19 (s, 3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ = 160.2, 148.7, 147.5, 138.6, 27.8; IR (film): $\tilde{\upsilon}$ = 3270, 3225, 3080, 3025, 1700, 1568, 1510, 1427, 1395, 1233, 1107, 1021, 866, 702 cm^{-1}; MS (ESI): *m/z*: calcd for $C_5H_6N_4NaO_3$: 193.0, found: 193.0 $[M + Na]^+$.

6-Amino-5-nitroso-1-(phenylmethyl)-2,4(1H,3H)-pyrimidinedione (8)

Yield: 93%; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 13.33 (s, 1H), 11.66 (s, 1H), 9.12 (s, 1H), 7.24-7.36 (m, 5H), 5.08 (s, 2H); ^{13}C NMR (75 MHz, DMSO-d_6): δ = 160.4, 148.9, 146.9, 138.6, 134.6, 128.6, 127.4, 126.3, 43.3; IR (film): $\tilde{\upsilon}$ = 3145, 2969, 2773, 1729, 1691, 1636, 1509, 1448, 1410, 1247, 1161, 725 cm^{-1}; MS (ESI): *m/z*: calcd for $C_{11}H_{10}N_4NaO_3$: 269.1, found: 269.0 $[M + Na]^+$.

General procedure for the preparation of 1-alkyl-5,6-diaminouracil (9, 10)

To a stirred mixture containing 1-alkyl-5-nitroso-4-aminouracil (1 eq) in aqueous ammonia (25%, 0.6 M), sodium dithionite (3.5 eq) was added in small portions. The temperature raised to 35 °C and the purple color disappeared gradually. When no more temperature increase was observed, the reaction mixture was stirred at 60 °C for an additional hour and then 17 h at room temperature in order to complete the reaction (total disappearance of the purple color). The solid was filtered off, thoroughly washed with cold water, and dried under high vacuum to obtain the corresponding products in pure form.

5,6-Diamino-1-methyl-2,4(1H,3H)-pyrimidinedione (9)

Yield: 78%; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 10.53 (s, 1H), 6.12 (s, 2H), 3.21 (s, 3H), 2.95 (s, 2H); IR (film): $\tilde{\upsilon}$ = 3351, 3313, 3148, 1667, 1636, 1579, 1509, 962, 712 cm^{-1}; MS (ESI): *m/z*: calcd for $C_5H_9N_4O_2$: 157.1, found: 157.0 $[M + H]^+$.

5,6-Diamino-1-(phenylmethyl)-2,4(1H,3H)-pyrimidinedione (10)

Yield: 90%; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 9.88 (s, 1H), 7.19-7.35 (m, 5H), 6.09 (s, 2H), 5.07 (s, 2H), 2.92 (s, 2H); ^{13}C NMR (75 MHz, DMSO-d_6): δ = 159.6, 149.6, 145.3, 136.8, 128.3, 127.0, 126.4, 96.4, 43.9; IR (film): $\tilde{\upsilon}$ = 3406, 3305, 3183, 2968, 2776, 1686, 1628, 1574, 1496, 1402, 942, 717 cm^{-1}; MS (ESI): *m/z*: calcd for $C_{11}H_{12}N_4NaO_2$: 255.1, found: 255.1 $[M + Na]^+$.

3,9-Dihydro-3-methyl-1H-purine-2,6-dione (11)

A solution of 1-methyl-5,6-diaminouracil (18.8 g, 120 mmol) in formic acid (6.2 mL) and water (140 mL) was heated to reflux for 3 h. After cooling to 20 °C, NaOH (9.6 g, 240 mmol) in water (12 mL) was added, and the mixture was heated to reflux for 1 h. After cooling to 0 °C, the mixture was acidified by addition of acetic acid. The precipitate was filtered, washed with water, and dried to afford 11 as a light yellow solid (15.5 g, 78% yield). ^{1}H NMR (300 MHz, DMSO-d_6): δ = 11.09 (s, 1H), 8.04 (s, 1H), 3.37 (s, 3H), N_7H not observed; ^{13}C NMR (100 MHz, DMSO-d_6): δ = 154.6, 151.1, 149.2, 140.2, 106.9, 28.8; IR (film): $\tilde{\upsilon}$ = 3371, 3102, 3011, 1689, 1401, 1164, 827, 742 cm^{-1}; MS (ESI): *m/z*: calcd for $C_6H_7N_4O_2$: 167.1, found: 167.0 $[M + H]^+$.

3,9-Dihydro-3-(phenylmethyl)-1H-purine-2,6-dione (12)

A solution of 1-benzyl-5,6-diaminouracil (761 mg, 3.28 mmol) in formic acid (15 mL) was heated to reflux for 1 h and then the formic acid was evaporated under reduced pressure. The residue was dissolved in

aqueous NaOH (10%, 15 mL) and ethanol (5 mL). The mixture was heated to reflux for 1 h, and then concentrated. After cooling to 0 °C, the mixture was acidified by addition of concentrated HCl. The precipitate was filtered, washed with water, and dried to afford 12 as a light yellow solid (714 mg, 90% yield). ^{1}H NMR (400 MHz, DMSO-d_6): δ = 13.54 (s, 1H), 11.18 (s, 1H), 8.01 (s, 1H), 7.23-7.32 (m, 5H), 5.11 (s, 2H); ^{13}C NMR (75 MHz, DMSO-d_6): δ = 154.6, 150.9, 149.1, 140.6, 137.0, 128.3, 127.4, 127.1, 107.0, 44.9; IR (film): $\tilde{\upsilon}$ = 3145, 2986, 2906, 1662, 1551, 1414, 1260, 1065, 730 cm^{-1}; MS (ESI): *m/z*: calcd for $C_{12}H_{10}N_4NaO_2$: 265.1, found: 265.0 $[M + Na]^+$.

General procedure for the bromination of 3-alkylxanthines (13, 14)

To a stirring suspension of 3-alkylxanthine (1 eq) and sodium acetate (2 eq) in glacial acetic acid (0.39 M) was added bromine dropwise (1.2 eq). The mixture was stirred at 65 °C for 2 h. After cooling to room temperature the precipitate was filtered, washed with acetic acid, water, and dried under vacuum to give the corresponding 3-alkyl-8-bromoxanthine as beige powders in pure form.

8-Bromo-3,9-dihydro-3-methyl-1H-purine-2,6-dione (13)

Yield: 84%; ^{1}H NMR (300 MHz, DMSO-d_6): δ = 11.14 (s, 1H), 3.31 (s, 3H), N_7H not observed; ^{13}C NMR (75 MHz, DMSO-d_6): δ = 153.6, 150.7, 149.4, 124.1, 109.7, 28.8; IR (film): $\tilde{\upsilon}$ = 3577, 3362, 3029, 2968, 2832, 1667, 1468, 1268, 861, 748 cm^{-1}; MS (ESI): *m/z*: calcd for $C_6H_6BrN_4O_2$: 244.9, found: 244.8 $[M + H]^+$.

8-Bromo-3,9-dihydro-3-(phenylmethyl)-1H-purine-2,6-dione (14)

Yield: 73%; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 11.23 (s, 1H), 7.23-7.34 (m, 5H), 5.06 (s, 2H), N_7H not observed; ^{13}C NMR (100 MHz, DMSO-d_6): δ = 153.8, 150.6, 149.2, 136.8, 128.4, 127.3, 127.2, 124.7, 110.1, 45.1; IR (film): $\tilde{\upsilon}$ = 3027, 2850, 2817, 1702, 1667, 1541, 1460, 1368, 1340, 739 cm^{-1}; MS (ESI): *m/z*: calcd for $C_{12}H_{10}BrN_4O_2$: 320.9, found: 321.1 $[M + H]^+$.

2.7 Acknowledgment

We thank Marino Convertino and Dr. Peter Kolb for interesting discussions. The calculations were performed on the Matterhorn Beowulf cluster at the Informatikdienste of the University of Zurich. This work was supported in part by a grant of the Swiss National Science Foundation to A.C. and the Organic Chemistry Institute of the University of Zürich to C.N..

2.8 References

(1) Folkman, J. Angiogenesis in cancer, vascular, rheumatoid and other disease. *Nat. Med.* **1995**, *1*, 27-31.

(2) Folkman, J. Tumor angiogenesis: therapeutic implications. *N. Engl. J. Med.* **1971**, *285*, 1182-1186.

(3) Adams, R. H. Vascular patterning by Eph receptor tyrosine kinases and ephrins. *Semin. Cell. Dev. Biol.* **2002**, *13*, 55-60.

(4) Martiny-Baron, G.; Korff, T.; Schaffner, F.; Esser, N.; Eggstein, S.; Marme, D.; Augustin, H. G. Inhibition of tumor growth and angiogenesis by soluble EphB4. *Neoplasia* **2004**, *6*, 248-257.

(5) Kertesz, N.; Krasnoperov, V.; Reddy, R.; Leshanski, L.; Kumar, S. R.; Zozulya, S.; Gill, P. S. The soluble extracellular domain of EphB4 (sEphB4) antagonizes EphB4-EphrinB2 interaction, modulates angiogenesis, and inhibits tumor growth. *Blood* **2006**, *107*, 2330-2338.

(6) Kolb, P.; Kipouros, C. B.; Huang, D.; Caflisch, A. Structure-based tailoring of compound libraries for high-throughput screening: discovery of novel EphB4 kinase inhibitors. *Proteins* **2008**, *73*, 11-18.

(7) Miyazaki, Y.; Nakano, M.; Sato, H.; Truesdale, A. T.; Stuart, J. D.; Nartey, E. N.; Hightower, K. E.; Kane-Carson, L. Design and effective synthesis of novel templates, 3,7-diphenyl-4-amino-thieno and furo-[3,2-*c*]pyridines as protein kinase inhibitors and *in vitro* evaluation targeting angiogenetic kinases. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 250-254.

(8) Bardelle, C.; Cross, D.; Davenport, S.; Kettle, J. G.; Ko, E. J.; Leach, A. G.; Mortlock, A.; Read, J.; Roberts, N. J.; Robins, P.; Williams, E. J. Inhibitors of the tyrosine kinase EphB4. Part 1: Structure-based design and optimization of a series of 2,4-bis-anilinopyrimidines. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 2776-2780.

(9) Lombardo, L. J.; Lee, F. Y.; Chen, P.; Norris, D.; Barrish, J. C.; Behnia, K.; Castaneda, S.; Cornelius, L. A. M.; Das, J.; Doweyko, A. M.; Fairchild, C.; Hunt, J. T.; Inigo, I.; Johnston, K.; Kamath, A.; Kan, D.; Klei, H.; Marathe, P.; Pang, S. H.; Peterson, R.; Pitt, S.; Schieven, G. L.; Schmidt, R. J.; Tokarski, J.; Wen, M. L.; Wityak, J.; Borzilleri, R. M. Discovery of N-(2-chloro-6-methylphenyl)-2-(6-(4-(2-hydroxyethyl)-piperazin-1-yl)-2-methylpyrimidin-4-ylamino)thiazole-5-carboxamide (BMS-354825), a dual Src/Abl kinase inhibitor with potent antitumor activity in preclinical assays. *J. Med. Chem.* **2004**, *47*, 6658-6661.

(10) Bardelle, C.; Coleman, T.; Cross, D.; Davenport, S.; Kettle, J. G.; Ko, E. J.; Leach, A. G.; Mortlock, A.; Read, J.; Roberts, N. J.; Robins, P.; Williams, E. J. Inhibitors of the tyrosine kinase EphB4. Part 2: structure-based discovery and optimisation of 3,5-bis substituted anilinopyrimidines. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5717-5721.

(11) Karaman, M. W.; Herrgard, S.; Treiber, D. K.; Gallant, P.; Atteridge, C. E.; Campbell, B. T.; Chan, K. W.; Ciceri, P.; Davis, M. I.; Edeen, P. T.; Faraoni, R.; Floyd, M.; Hunt, J. P.; Lockhart, D. J.; Milanov, Z. V.; Morrison, M. J.; Pallares, G.; Patel, H. K.; Pritchard, S.; Wodicka, L. M.; Zarrinkar, P. P. A quantitative analysis of kinase inhibitor selectivity. *Nat. biotechnol.* **2008**, *26*, 127-132.

(12) McInnes, C. Virtual screening strategies in drug discovery. *Curr. Opin. Chem. Biol.* **2007**, *11*, 494-502.

(13) Huang, D.; Caflisch, A. Library screening by fragment-based docking. *J. Mol. Recognit.* **2010**, *23*, 183-193.

(14) Papesch, V.; Schroeder, E. F. Synthesis of 1-mono-substituted and 1,3-di-substituted 6-amino-uracils - diuretic activity. *J. Org. Chem.* **1951**, *16*, 1879-1890.

(15) Merlos, M.; Gomez, L.; Vericat, M. L.; Bartroli, J.; Garciarafanell, J.; Forn, J. Structure-activity-relationships in a series of xanthine derivatives with antibronchoconstrictory and bronchodilatory activities. *Eur. J. Med. Chem.* **1990**, *25*, 653-658.

(16) Baraldi, P. G.; Preti, D.; Tabrizi, M. A.; Fruttarolo, F.; Romagnoli, R.; Zaid, N. A.; Moorman, A. R.; Merighi, S.; Varani, K.; Borea, P. A. New pyrrolo[2,1-*f*]purine-2,4-dione and imidazo[2,1-*f*]purine-2,4-dione derivatives as potent and selective human A3 adenosine receptor antagonists. *J. Med. Chem.* **2005**, *48*, 4697-4701.

(17) Traxler, P.; Furet, P. Strategies toward the design of novel and selective protein tyrosine kinase inhibitors. *Pharmacol. Ther.* **1999**, *82*, 195-206.

(18) Bamborough, P.; Angell, R. M.; Bhamra, I.; Brown, D.; Bull, J.; Christopher, J. A.; Cooper, A. W.; Fazal, L. H.; Giordano, I.; Hind, L.; Patel, V. K.; Ranshaw, L. E.; Sims, M. J.; Skone, P. A.; Smith, K. J.; Vickerstaff, E.; Washington, M. N-4-Pyrimidinyl-1H-indazol-4-amine inhibitors of Lck: indazoles as phenol isosteres with improved pharmacokinetics. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 4363-4368.

(19) Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The protein kinase complement of the human genome. *Science* **2002**, *298*, 1912-1934.

(20) Huang, D.; Zhou, T.; Lafleur, K.; Nevado, C.; Caflisch, A. Kinases selectivity potential for inhibitors targeting the ATP binding site: a network analysis. *Bioinformatics* **2010**, *26*, 198-204.

(21) Bamborough, P.; Drewry, D.; Harper, G.; Smith, G. K.; Schneider, K. Assessment of chemical coverage of kinome space and its implications for kinase drug discovery. *J. Med. Chem.* **2008**, *51*, 7898-7914.

(22) Sun, L.; Liang, C.; Shirazian, S.; Zhou, Y.; Miller, T.; Cui, J.; Fukuda, J. Y.; Chu, J. Y.; Nematalla, A.; Wang, X. Y.; Chen, H.; Sistla, A.; Luu, T. C.; Tang, F.; Wei, J.; Tang, C. Discovery of 5-[5-Fluoro-2-oxo-1,2-dihydroindol-(3Z)-ylidenemethyl]-2,4-dimethyl-1H-pyrrole-3-carboxylic acid (2-diethylaminoethyl)amide, a novel tyrosine kinase inhibitor targeting vascular endothelial and platelet-derived growth factor receptor tyrosine kinase. *J. Med. Chem.* **2003**, *46*, 1116-1119.

(23) Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today* **2004**, *9*, 430-431.

(24) Huang, D.; Luthi, U.; Kolb, P.; Cecchini, M.; Barberis, A.; Caflisch, A. *In silico* discovery of beta-secretase inhibitors. *J. Am. Chem. Soc.* **2006**, *128*, 5436-5443.

(25) Kolb, P.; Caflisch, A. Automatic and efficient decomposition of two-dimensional structures of small molecules for fragment-based high-throughput docking. *J. Med. Chem.* **2006**, *49*, 7384-7392.

(26) Scarsi, M.; Apostolakis, J.; Caflisch, A. Continuum electrostatic energies of macromolecules in aqueous solutions. *J. Phys. Chem. A* **1997**, *101*, 8098-8106.

(27) Caflisch, A.; Fischer, S.; Karplus, M. Docking by Monte Carlo minimization with a solvation correction: Application to an FKBP-substrate complex. *J. Comput. Chem.* **1997**, *18*, 723-743.

(28) Majeux, N.; Scarsi, M.; Apostolakis, J.; Ehrhardt, C.; Caflisch, A. Exhaustive docking of molecular fragments with electrostatic solvation. *Proteins* **1999**, *37*, 88-105.

(29) Majeux, N.; Scarsi, M.; Caflisch, A. Efficient electrostatic solvation model for protein-fragment docking. *Proteins* **2001**, *42*, 256-268.

(30) Budin, N.; Majeux, N.; Caflisch, A. Fragment-Based flexible ligand docking by evolutionary optimization. *Biol. Chem.* **2001**, *382*, 1365-1372.

(31) Cecchini, M.; Kolb, P.; Majeux, N.; Caflisch, A. Automated docking of highly flexible ligands by genetic algorithms: a critical assessment. *J. Comput. Chem.* **2004**, *25*, 412-422.

(32) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. Charmm - a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187-217.

(33) Brooks, B. R.; Brooks, C. L., 3rd; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **2009**, *30*, 1545-1614.

(34) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586-3616.

(35) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926-935.

(36) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33-38.

(37) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781-1802.

(38) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh ewald method. *J. Chem. Phys.* **1995**, *103*, 8577-8593.

(39) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant-pressure molecular-dynamics algorithms. *J. Chem. Phys.* **1994**, *101*, 4177-4189.

(40) Feller, S. E.; Zhang, Y. H.; Pastor, R. W.; Brooks, B. R. Constant-Pressure Molecular-Dynamics Simulation - the Langevin Piston Method. *J. Chem. Phys.* **1995**, *103*, 4613-4621.

(41) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical-integration of cartesian equations of motion of a system with constraints - molecular-dynamics of *N*-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327-341.

(42) Ballester, P. J.; Richards, W. G. Ultrafast shape recognition to search compound databases for similar molecular shapes. *J. Comput. Chem.* **2007**, *28*, 1711-1723.

(43) Granovsky, A. A. PC GAMESS/Firefly version 7.1.F, www http://classic.chem.msu.su/gran/gamess/index.html.

(44) Dumas, J.; Phillips, B.; Zhang, C.; Ladouceur, G. H.; Zhao, Q.; Hentemann, M.; Verma, S. K.; Zhu, Q.; Lavoie, R. C.; Fan, J. Benzofuran derivatives useful for treating hyper-proliferative disorders. WO/2005/014566, **2005**.

Chapter 3: Synthesis and Biological Evaluation of a Second Generation of EphB4 Inhibitors

3 Synthesis and Biological Evaluation of a Second Generation of EphB4 Inhibitors

3.1 Introduction

Fragment based docking on tridimensional known structures is a computational tool frequently used in *in silico* drug discovery. Starting from a library of about 700,000 compounds, a computational strategy was designed to generate a focused library of 21,418 molecules with promising anchor fragments for the ATP binding site of EphB4. Automatic docking and subsequent in vitro evaluations yielded compounds **1** and **2** which inhibit EphB4 activity in the low micromolar range.[1] Their predicted binding mode involves two hydrogen bonds between the pyrimidine ring and the hinge region of the ATP binding pocket (Figure 1, left, interactions with Glu694 and Met696 are indicated in dotted lines). With these results in hand, a computational aided medicinal chemistry campaign was implemented to optimize the binding affinity of these compounds.[2] In fact, addition of a methyl and a hydroxyl group on the phenyl ring provided compound **3** (Figure 1, right), which showed single digit nanomolar activity against EphB4. Analysis of a 45-ns MD simulation suggested that the hydrogen bonds formed between the pyrimidine ring of compound **3** and the hinge region are conserved during the entire run. During the first 12 ns, the hydroxyl group is involved in two additional hydrogen bonds with the side chain of Glu664 and the backbone NH of Asp758. A reorientation of these residues after 12 ns induces a new hydrogen bonding pattern, where the hydroxyl group of compound **3** interacts with the backbone carbonyl of Ser757. Finally, the methyl group is involved in favorable van der Waals interactions within the hydrophobic pocket, which, taken together with the additional hydrogen bonds formed between the molecule and the ATP binding site, explains the higher potency of compound **3** vs. compounds **1** and **2**.

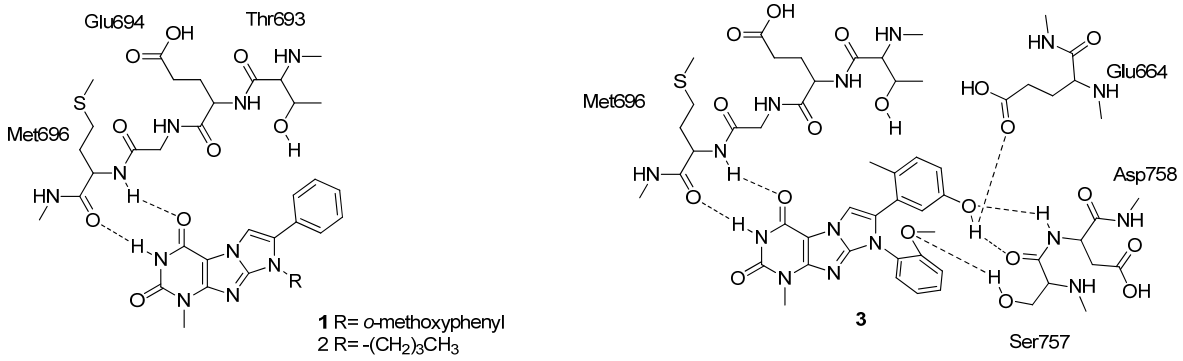


Figure 1. Binding mode of compounds **1**, **2** (left) and **3** (right)

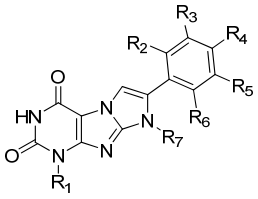
Further biological studies on compound **3** showed that its pharmacological properties could be improved. Significant discrepancies between enzymatic and cellular based assay were observed, which

could be due to lack of permeability or substrate recognition by efflux transporters. In addition, phenols usually undergo glucuronidation during phase II metabolism, thus have poor pharmacokinetic properties. Herein, we present a multidisciplinary approach aimed at the development of a second generation of EphB4 inhibitors with high enzymatic and cellular affinity. In addition, the binding modes of our lead compounds have been confirmed by X-ray crystallography.

3.2 Lead optimization strategy

The inhibition activity of compounds **1** and **2** on EphB4-transfected Chinese hamster ovary (CHO) cells has been shown to be relatively low.[1] However, as compound **3** inhibited EphB4 enzymatic activity in the low nanomolar range, we decided to explore its potency on cellular based assays. Thus, cellular phosphorylation assays were done at Proqinase on murine embryonal fibroblast cells (MEF) transfected with myc-tagged human EphB4. After incubation with the corresponding inhibitor and stimulation with ephrinB2-Fc, autophosphorylation of EphB4 was quantified via sandwich ELISA. While compound **3** showed single digit nanomolar activity in enzymatic assay, its cellular activity was reduced by almost two orders of magnitude, with IC_{50} values higher than 100 nM. The enzymatic and cellular activities of compounds **1**, **2** and **3** are summarized in Table 1.

Table 1. EphB4 inhibition data of xanthine derivatives.



Compound	R_1	R_2	R_3	R_4	R_5	R_6	R_7	Enzymatic IC_{50} (nM)[a]	Cell IC_{50} (nM)
1	Me	H	H	H	H	H	*o*-methoxyphenyl	3300 (4350)	14% at 20 µM
2	Me	H	H	H	H	H	butyl	1900 (538)	neg.
3	Me	Me	H	H	OH	H	*o*-methoxyphenyl	5 (1.6)	130

[a] IC_{50} values determined at Reaction Biology are given in parentheses.

Discrepancies between enzymatic and cellular activity (5 nM in enzymatic assay vs. 130 nM in cellular phosphorylation assay for compound **3**) prompted us to evaluate further pharmacological properties, such as cell permeability on Caco-2 monolayers and active efflux by cell transporters (Absorption Systems). Caco-2 cells express many transport proteins present on epithelial cells of the small intestine, and form tight junctions between each other. Thus, measurements of permeability through Caco-2

monolayers may predict absorption across intestinal tissues *in vivo*. Interestingly, compounds **1** and **2** bearing hydrogens at position R_2 and R_5 showed high cell permeability and no significant efflux was observed (Table 2), the latter showing the highest membrane permeation. As suggested by the cellular phosphorylation assays, compound **3** with a methyl group at R_2 and a hydroxyl group at R_5 showed high permeability although it was efficiently effluxed by P-glycoprotein transporters.

Table 2. Permeability measurements through Caco-2 cells.

Compound ID	%Recovery		Papp ($x10^{-6}$ cm/s)		Efflux ratio*	Permeability classification	Significant efflux
	A-B	B-A	A-B	B-A			
1	71	76	32.3	35.8	1.1	High	No
2	69	78	42.5	42.5	1.0	High	No
3	51	74	3.13	23.0	7.3	High	Yes

Permeability Classification: (Papp A -> B) < 1.0 X 10^{-6}cm/s: **Low**
(Papp A -> B) > 1.0 X 10^{-6} cm/s: **High**
Significant Efflux: Efflux > 3.0 and (Papp B -> A) > 1.0 X 10^{-6} cm/s
*Efflux ratio = (Papp B -> A) / (Papp A -> B)

As phenols are known to be detrimental for the pharmacokinetic properties of drugs, we decided to avoid the presence of a hydroxyl group at position R_5. According to the permeability measurements, replacement of the anisidine side chain by an alkyl chain appeared to be promising (compare compounds **1** with Papp = 32.3-35.8 x 10^{-6} cm/s and **2** with Papp = 42.5 x 10^{-6} cm/s in Table 2). Thus, we investigated the synthesis of molecules bearing a methyl group at position R_2 and a butyl, pentyl or hexyl chain at position R_7. In addition, we decided to introduce a fluorine at position R_4, since such a substitution proved to be very useful in terms of metabolic stability and lipophilicity.[3] In order to form a halogen bond with Glu664,[4] we also examined the addition of a chlorine at position R_5. Finally, we decided to introduce a chlorine atom or a trifluoromethyl group at position R_2 to improve van der Waals interactions with the side chains of Val629, Ala645, and Thr693.

Rigidification of biologically active molecules has been shown to be beneficial for selectivity properties,[5] as the number of accessible conformations able to interact with secondary targets is restricted. In addition, ortho-substitution of bicyclic molecules with methyl or methoxy groups has been shown to decrease crystal packing and improve aqueous solubility.[6] Therefore, we investigated the synthesis of a bis-ortho-methyl substitution. Finally, poor oral bioavailability of the ortho-methyl/meta-hydroxyl pattern has been published for other kinase inhibitors,[5,7,8] leading to the replacement of the

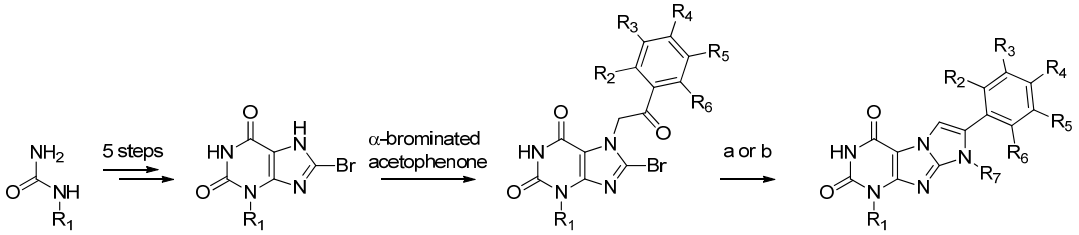
phenol ring by an indazole moiety.[5] This transformation was thought to give access to a molecule with a similar hydrogen bonding pattern than compound **3**.

With these designs in mind, we set out to explore the synthesis of these customized analogues in order to later evaluate their pharmacological profile.

3.3 Synthesis

Recently, there has been growing interest in the synthesis of imidazoxanthines, as a few derivatives proved to be potent serotonin,[9] A_3 adenosine[10,11] or kinase receptor antagonists.[2] Following our previously developed route, the synthesis of the key intermediate 8-bromoxanthine was achieved in five steps starting from the corresponding alkylurea (Scheme 1). After alkylation with α-brominated acetophenones, cyclization usually occurs by refluxing the resulting intermediate in ethanol in the presence of a primary amine.[2,9,10] In addition, Baraldi and coworkers reported a cyclization using Mitsonobu reaction conditions.[11] However, to the best of our knowledge, the synthesis of bis-ortho substituted imidazoxanthines has not been described yet. In addition, none of the mono-ortho substituted molecules reported so far has an alkyl chain at position R_7.

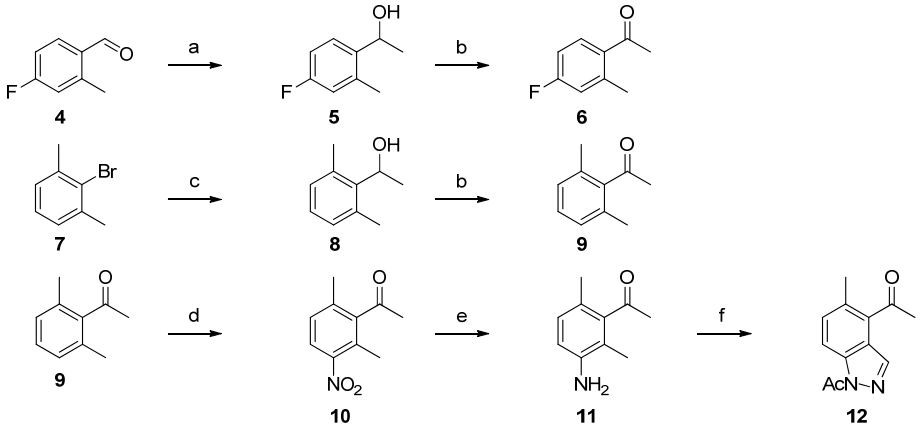
Scheme 1[a]



[a]Reagents and conditions: (a) Primary amine, EtOH or methoxyethanol, reflux, 15 h. (b) Primary amine, PPh_3, DEAD, THF, 25 °C, 1 h.

The synthesis of non-commercially available acetophenones has been summarized in Scheme 2. Treatment of commercially available 4-fluoro-2-methylbenzaldehyde **4** with methyllithium followed by oxidation with PCC provided methyl ketone **6**. The synthesis of 1-(2,6-dimethylphenyl)ethanone started with 2-bromo-1,3-dimethylbenzene **7**, which after Br/Li exchange in the presence of Buli, reacted with acetaldehyde to afford the secondary alcohol **8**. Subsequent oxidation using PCC gave the expected acetophenone **9**,[12] which was also used in the synthesis of 1,1'-(5-methyl-1H-indazole-1,4-diyl)diethanone **12**. Indeed, reaction with KNO_3 selectively nitrated compound **9** at the meta position.[13] Next, reduction in the presence of iron and cyclization using isoamylnitrite afforded the protected indazole derivative **12**.[14]

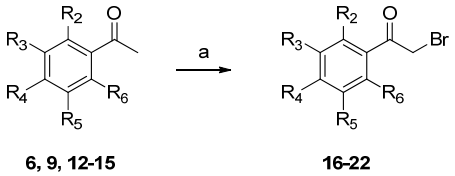
Scheme 2[a]



[a]Reagents and conditions: (a) MeLi, THF, 25 °C, 1 h. (b) PCC, CH_2Cl_2, 25 °C, 1 h. (c) BuLi, Acetaldehyde, THF, -78 °C, 1.5 h. (d) KNO_3, H_2SO_4, 0 °C, 1 h. (e) Fe, AcOH, EtOH, 90 °C, 2 h. (f) NaOAc, Ac_2O, $CHCl_3$, 25 °C, 30 min, then iAmONO, 60 °C, 7 h.

α-Bromination of commercially available and synthesized acetophenones was achieved in the presence of copper(II) bromide in chloroform according to Scheme 3.[15]

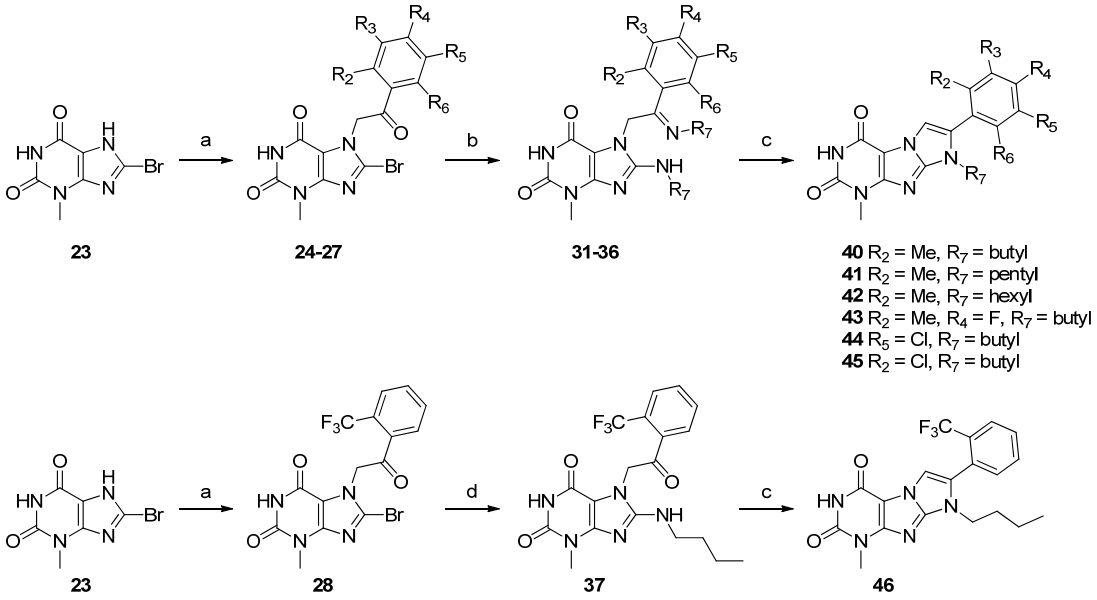
Scheme 3[a]



[a]Reagents and conditions: (a) $CuBr_2$, $CHCl_3$, EtOAc, reflux, 15 h (21-64% of yield).

Alkylation at the N_7 position of the xanthine core with meta-substituted or mono-ortho-substituted α-haloketones was accomplished using *N,N*-diisopropylethylamine in DMF (Scheme 4). However, cyclization with alkyl amines under standard conditions was never observed. Instead, intermediates **31**-**36** could be isolated and further transformed upon reaction with aluminium chloride to afford the expected products **40**-**45**. This method was used as well to generate compound **44**, as a prolonged heating in the presence of butylamine could lead to an additional substitution of the chlorine atom. Interestingly, as position R_2 was substituted by a trifluoromethyl group, reaction with butylamine gave intermediate **37** which could be converted into **46** upon treatment with $AlCl_3$.

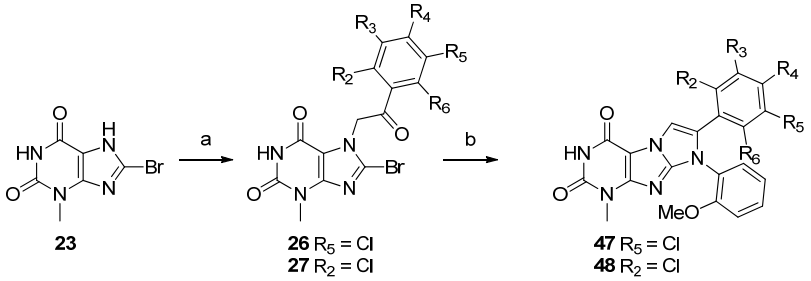
Scheme 4[a,b]



[a]Reagents and conditions: (a) α-haloketone, DIPEA, DMF, 25 °C, 17 h. (b) Primary amine, sealed tube, 180 °C, 30 min. (c) $AlCl_3$, EtOH, sealed tube, 180 °C, 2-15 h. (d) Butylamine, EtOH, sealed tube, 180 °C, 1 h.
[b]Unless otherwise stated, R_n = H.

Cyclization in the presence of anisidine was done as previously reported,[2] giving final products **47** and **48** (Scheme 5).

Scheme 5[a,b]

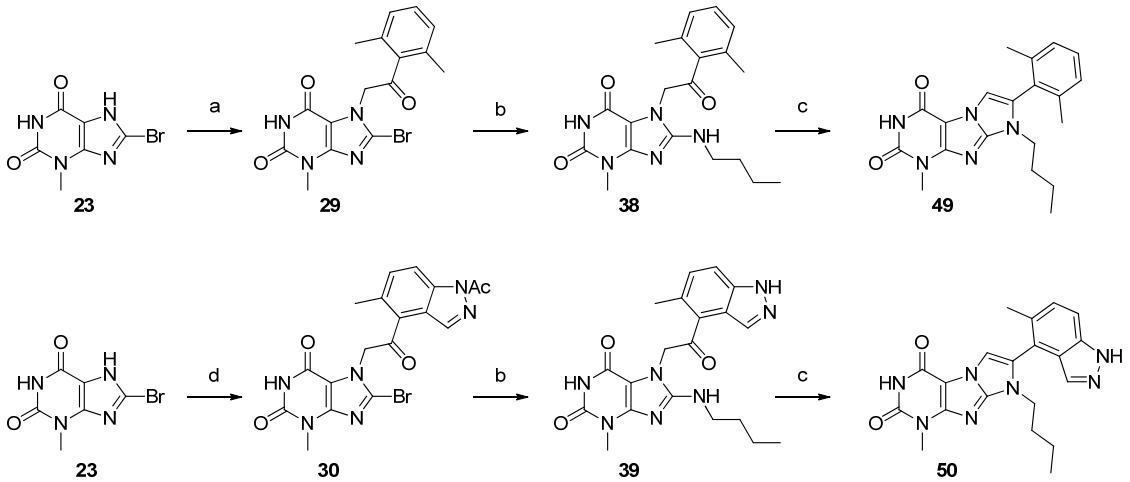


[a]Reagents and conditions: (a) α-haloketone, DIPEA, DMF, 25 °C, 17 h. (b) Anisidine, EtOH, sealed tube, reflux, 15 h.
[b]Unless otherwise stated, R_n = H.

Alkylation of **23** with bis-ortho-substituted α-haloketones required heating (Scheme 6), and subsequent cyclization in the presence of anisidine failed. Reaction with butyl amine provided

intermediates **38** and **39**, which could finally be cyclized by heating in a concentrated solution of boron trifluoride diethyletherate (compounds **49** and **50**).

Scheme 6[a]



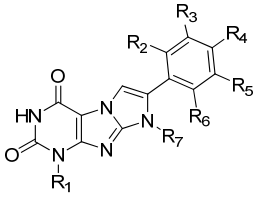
[a]Reagents and conditions: (a) 2-bromo-1-(2,6-dimethylphenyl)ethanone, DIPEA, DMF, 110 °C, 2 h. (b) Butylamine, EtOH, sealed tube, 180 °C, 0.5-2 h. (c) $BF_3 \cdot OEt_2$, DCM, sealed tube, 180 °C, 0.5-6 h. (d) KOH, EtOH, reflux, 2 h, then 1-(1-acetyl-5-methyl-1H-indazol-4-yl)-2-bromoethanone, DMF, 110 °C, 2 h.

3.4 Enzymatic and cellular based assay

The inhibition activity of our second generation of EphB4 inhibitors was first evaluated on enzymatic assays (Table 3). Removal of the hydroxyl group at R_5 and replacement of the anisidine chain by a butyl chain yielded **40** with an IC_{50} around 50 nM. While an elongation of the butyl chain (compounds **41** and **42**) didn't improve the affinity, the introduction of a fluorine at position R_4 yielded compound **43** with similar binding affinity than compound **40**. However, addition of a chlorine at position R_5 (compounds **44** and **47**) as well as replacement of the methyl group by a chlorine atom or a trifluoromethyl moiety (compounds **45**, **46** and **48**) was detrimental for binding affinity. Finally, compound **49** bearing methyl groups at position R_2 and R_6 showed micromolar activity, whereas compound **50** with an indazole ring had a similar inhibition activity than our original compound **3**.

Thus, together with compounds **40**, **41**, **43**, **45** and **48**, cellular phosphorylation assays have been performed (Proqinase). In fact, except for compound **48** which had a micromolar affinity, every tested compound showed comparable levels of activity to the ones obtained in the enzymatic assays. Interestingly, compounds **40** and **43** showed IC_{50} values below 100 nM, thus being the most promising molecules of this second optimization campaign.

Table 3. EphB4 inhibition data of xanthine derivatives.



Compound	R_1	R_2	R_3	R_4	R_5	R_6	R_7	Enzymatic IC_{50} (nM)[a]	Cell IC_{50} (nM)
40	Me	Me	H	H	H	H	butyl	56 (82)	50
41	Me	Me	H	H	H	H	pentyl	177	230
42	Me	Me	H	H	H	H	hexyl	1300	n.d.
43	Me	Me	H	F	H	H	butyl	91 (139)	68
44	Me	H	H	H	Cl	H	butyl	22% @ 10uM	n.d.
45	Me	Cl	H	H	H	H	butyl	182	270
46	Me	CF_3	H	H	H	H	butyl	15% @ 10uM	n.d.
47	Me	H	H	H	Cl	H	*o*-methoxyphenyl	12% @ 10uM	n.d.
48	Me	Cl	H	H	H	H	*o*-methoxyphenyl	300	2200
49	Me	Me	H	H	H	Me	butyl	49% @ 10uM	n.d.
50	Me	Me	H	H	Indazole		butyl	14 (133)	150

[a] IC_{50} values determined at Reaction Biology are given in parentheses.

In addition, the cell permeation of compounds **40** and **50** through Caco-2 monolayers was further investigated (Absorption Systems). As predicted by cellular phosphorylation assays, both compounds had a high permeability. However, only compound **40** was not effluxed by transport proteins (Table 4).

Table 4. Permeability measurements through Caco-2 cells.

Compound ID	%Recovery		Papp ($x10^{-6}$ cm/s)		Efflux ratio*	Permeability classification	Significant efflux
	A-B	B-A	A-B	B-A			
40	85	89	42.7	28.1	0.7	High	No
50	84	96	8.02	49.0	6.1	High	Yes

Permeability Classification: (Papp A -> B) < 1.0 X 10^{-6}cm/s: **Low**
(Papp A -> B) > 1.0 X 10^{-6} cm/s: **High**
Significant Efflux: Efflux > 3.0 and (Papp B -> A) > 1.0 X 10^{-6} cm/s
*Efflux ratio = (Papp B -> A) / (Papp A -> B)

3.5 Selectivity profile

In order to assess the selectivity profile of compound **40**, enzymatic assays were performed on a panel of 124 kinases (National Centre for Protein Kinase Profiling at the University of Dundee). Out of these 124 kinases, only six of them (Src, Lck, EphA2, EphA4, EphB1 and EphB2) were very strongly inhibited. In addition, four other kinases (RIPK2, Yes-1, EphB3 and EphB4) showed relatively strong inhibition by compound **40** (see section 3.9). It is important to note that all of them have a threonine as a gatekeeper. Overall, compound **40** has a similar selectivity profile than **3**, with a strong inhibition of a relatively small fraction of the human kinome (Figure 2).

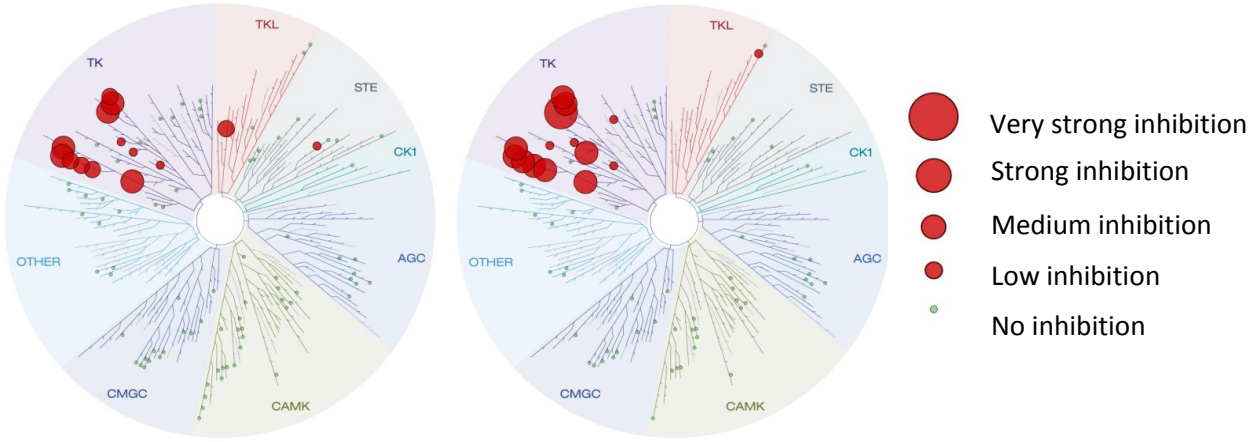


Figure 2. Selectivity profile of compound **40** (left) and **3** (right). Affinity in enzymatic assays is defined as very strong, strong, medium and low (corresponding to <10%, 10-20%, 20-30%, and 30-70% of remaining activity with respect to a DMSO control).

In addition, the inhibition activity of our most potent compounds **3**, **40**, **43** and **50** has been assessed on Abl, Src, Lck and Yes (Reaction Biology Corporation, Table 5). While compound **3** showed single digit nanomolar activity on Abl, Src and Lck, compounds **40**, **43** and **50** only retained high inhibition on Src.

Table 5. Abl, Lck, Src and Yes1 inhibition data of xanthine derivatives.

Compound ID	Abl (nM)	Src (nM)	Lck (nM)	Yes1 (nM)
3	0.92	1.14	1.3	n.d.
40	578	71	493	284
43	1282	110	617	496
50	203	46	264	160.4

3.6 Confirmation of the binding mode by X-ray crystallography

The binding modes of compounds **3**, **40** and **50** were further investigated by X-ray crystallography (performed by Dr. Jing Dong, Table 6). As compound **3** showed low nanomolar activity on other Eph kinases, we decided to focus on the crystallization of EphA3, whose structure resolution seemed higher than EphB4. Thus, the most important parameters obtained for the crystallization of **3**, **40** and **50** bound to EphA3 are summarized in Table 6 (hanging drop vapor diffusion method, see section 3.8.6).

Table 6. X-ray data collection.

	3	**40**	**50**
Space group	P 1 21 1	P 1 21 1	P 1 21 1
Unit cell			
a (Å)	52.6	53.29	53.06
b (Å)	38.2	38.25	38.18
c (Å)	75.7	76.04	75.84
Resolution range (Å)	38.2-2.2	34.0-1.9	38.2-2.1
Unique reflections	15076 (2090)	23747 (3297)	17532 (2402)
<I/σ(I)>	8.8 (3.0)	8.9 (2.8)	10.6 (3.9)
R merge	0.103 (0.457)	0.084 (0.405)	0.086 (0.334)
Completeness (%)	98.4 (94.6)	98.5 (94.8)	98.8 (93.4)
Multiplicity	3.7 (3.5)	3.0 (2.9)	3.5 (3.4)
Refinement			
Resolution range (Å)	30.68-2.20	29.79-1.90	29.70-2.10
R factor / R free	19.15 / 23.25	18.45 / 22.34	17.37-22.53
Mean B factors ($Å^2$)	30.80	24.30	28.80
RMS bonds (Å)	0.0077	0.0071	0.0071
RMS angles (°)	1.505	1.377	1.365

Analysis of the structural features observed by X-ray crystallography revealed that the most important interactions were already predicted by MD simulation on compound **3** (Figure 3). The pyrimidine ring is involved in two hydrogen bonds with the backbone CO and NH of Met696, and the methoxy group of the anisidine side chain interacts with the hydroxyl group of Ser757. Importantly, the

hydroxyl group of the phenol moiety is involved in hydrogen bond interactions with the backbone NH of Asp758 and the carboxylate group of Glu664.

A reorientation of the peptide bond linking Ser757 and Asp758 was predicted by MD simulation, leading to the formation of a hydrogen bond between the phenol moiety of **3** and the carbonyl group of Ser757. In contrast, the hydroxyl group only interacts with Glu664 and Asp758 in the crystal structure. The observed rotation of the peptide bond during the simulation might be due to the absence of Mg^{2+} which stabilizes the orientation of Asp758.

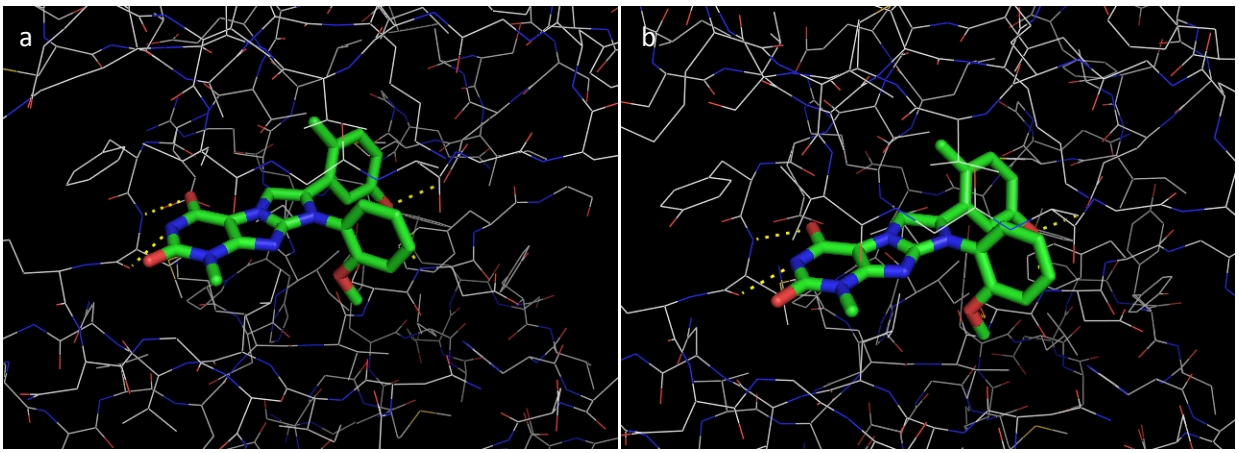


Figure 3. (a) Binding mode of compound **3** obtained by automatic docking on EphB4. (b) Binding mode of compound **3** obtained by X-ray crystallography on EphA3.

The crystal structures of compounds **40** and **50** show that the key features of the binding mode are conserved, such as the hydrogen bond interactions between the pyrimidine ring and Met696, as well as the orientation of the phenyl ring in the deep ATP back pocket (Figure 4). Interestingly, the indazole ring of compound **50** is involved in two hydrogen bond interactions. While the NH group interacts with Glu664, the imine nitrogen is involved in a hydrogen bond with Asp758.

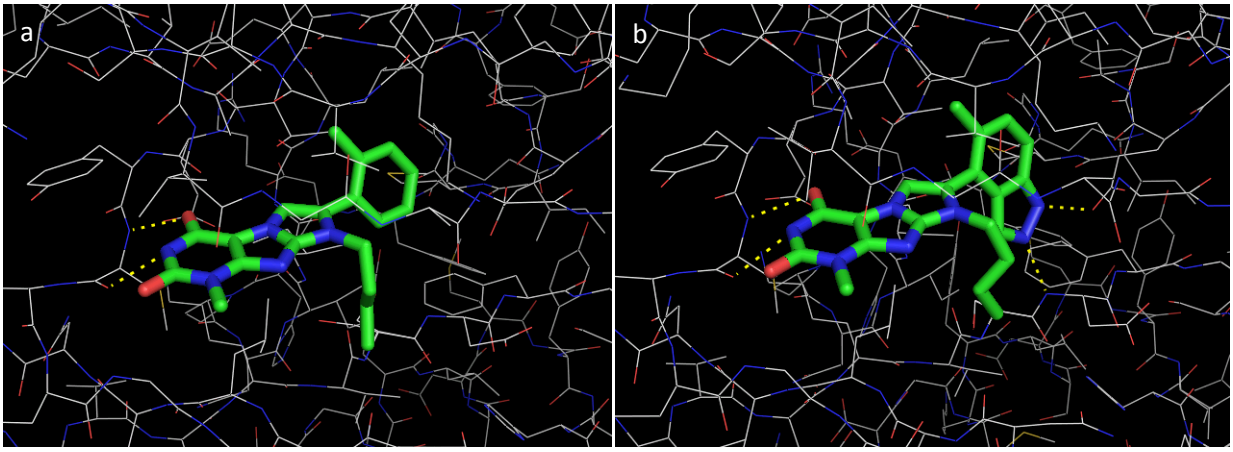


Figure 4. Binding mode of compounds **40** (a) and **50** (b) obtained by X-ray crystallography on EphA3.

3.7 Conclusions

Recently, a medicinal chemistry campaign supported by computational docking allowed the discovery of compound **3**, a single-digit nanomolar inhibitor of EphB4. However, further biological evaluation of compound **3** revealed a major discrepancy between enzymatic and cellular activity (5 nM in enzymatic assay vs. 130 nM in cellular phosphorylation assay). Furthermore, efficient efflux by P-glycoprotein transporters was observed on caco-2 monolayers.

Therefore, a second generation of EphB4 inhibitors was designed and synthesized. Compound **40** showed a 50 nM inhibition activity both in enzymatic and cellular based assays. Further biological analysis suggested that compound **40** was not effluxed on caco-2 cells, thus having more promising pharmacokinetic properties. In addition, the selectivity profile of **40** is similar to that of compound **3**, as strong inhibition activity was observed on only 5 percent of a 124 kinase panel.

Remarkably, the binding mode of compounds **3**, **40** and **50** suggested by docking and MD simulations in explicit solvent has been confirmed by X-ray crystallography. In particular, the important features of the predicted binding mode including the hydrogen bonding interactions and the orientation of the phenyl ring in the hydrophobic pocket were in agreement with the experimental results.

In summary, further synthetic efforts and additional biological analysis provided new promising compounds for *in vivo* studies.

3.8 Experimental section

3.8.1 Chemistry

All reactions, unless otherwise stated, were carried out under a nitrogen atmosphere using standard Schlenk-techniques. All reagents were used as received unless otherwise noted. Solvents were purchased in the best quality available, degassed by purging thoroughly with nitrogen and dried over activated molecular sieves of appropriate size. Alternatively, they were purged with argon and passed through alumina columns in a solvent purification system (Innovative Technology). Reactions were monitored by thin layer chromatography (TLC) using Merck TLC silica gel 60 F254. Flash column chromatography was performed over silica gel (230-400 mesh). NMR spectra were recorded on AV2 400 or AV2 500 MHz Bruker spectrometers. Chemical shifts are given in ppm. The spectra are calibrated to the residual ^{1}H and ^{13}C signals of the solvents. Multiplicities are abbreviated as follows: singlet (s), doublet (d), triplet (t), quartet (q), doublet-doublet (dd), quintet (quint), septet (sept), multiplet (m), and broad (br). Melting points were determined on a Büchi Melting Point B-540 instrument. 4-Fluoro-2-methylbenzaldehyde (**4**), 2-bromo-1,3-dimethylbenzene (**7**), 2′-chloroacetophenone (**13**), 3′-

chloroacetophenone (**14**) and 2′-(trifluoromethyl)acetophenone (**15**) were purchased from Fluka. 2-bromo-1-(o-tolyl)ethanone (**22**) was purchased from Synchem.

High-resolution electrospray ionization mass spectrometry was performed on a Finnigan MAT 900 (Thermo Finnigan, San Jose, CA, USA) double-focusing magnetic sector mass spectrometer. Ten spectra were acquired. A mass accuracy ≤2 ppm was obtained in the peak matching acquisition mode by using a solution containing 2 µL PEG200, 2 µL PPG450, and 1.5 mg NaOAc (all obtained from Sigma-Aldrich, Buchs, Switzerland) dissolved in 100 mL MeOH (HPLC Supra grade, Scharlau, E-Barcelona) as internal standard.

1-(4-Fluoro-2-methylphenyl)ethanol (5)

To a solution of 4-fluoro-2-methylbenzaldehyde (**4**, 0.8 mL, 6.62 mmol) in THF (20 mL) at -78 °C was added MeLi (1.6 M, 5.2 mL, 8.28 mmol) dropwise. The reaction was allowed to warm to room temperature and stirred for an additional hour. The reaction was quenched by addition of EtOAc, followed by water. The organic layer was washed with brine, dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The residue was immediately purified by column chromatography (gradient hexane:EtOAc 20:1 to 10:1) to afford the desired compound as a colorless oil (559 mg, 55% yield). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.46 (dd, *J* = 8.6 Hz, *J* = 5.9 Hz, 1H), 6.93-6.88 (m, 1H), 6.83 (dd, *J* = 9.7 Hz, *J* = 2.7 Hz, 1H), 5.08 (q, *J* = 6.4 Hz, 1H), 2.33 (s, 3H), 1.44 (d, *J* = 6.4 Hz, 3H), OH not observed; ^{13}C NMR (100 MHz, $CDCl_3$): δ = 161.7 (d, *J* = 244.7 Hz), 139.5 (d, *J* = 3 Hz), 136.6 (d, *J* = 7.7 Hz), 126.3 (d, *J* = 8.4 Hz), 116.8 (d, *J* = 20.9 Hz), 112.9 (d, *J* = 20.8 Hz), 66.4, 24.1, 18.9 (d, *J* = 1.4 Hz); IR (film): $\tilde{\upsilon}$ = 2975, 2929, 1681, 1580, 1354, 1234, 1115, 972, 865, 813 cm^{-1}; MS (GC/MS): *m/z*: calcd for $C_9H_{11}FO$: 154.1, found: 154.0.

1-(4-Fluoro-2-methylphenyl)ethanone (6)

1-(4-Fluoro-2-methylphenyl)ethanol (**5**, 559 mg, 3.62 mmol) was dissolved in CH_2Cl_2 (6 mL), and PCC (1.58 g, 7.33 mmol) was added to the mixture. After 1 h at 25 °C, the solution was filtered off over celite, concentrated under reduced pressure, and purified by flash chromatography on silica gel (gradient hexane:EtOAc 99:1 to 95:5) to afford a colorless oil (403 mg, 73% yield). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.74 (dd, *J* = 9.4 Hz, *J* = 5.8 Hz, 1H), 6.96-6.92 (m, 2H), 2.56 (s, 3H), 2.55 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 199.8, 164.1 (d, *J* = 253.2 Hz), 142.5 (d, *J* = 8.9 Hz), 133.7 (d, *J* = 2.9 Hz), 132.1 (d, *J* = 9.5 Hz), 118.9 (d, *J* = 21.2 Hz), 112.5 (d, *J* = 21.4 Hz), 29.4, 21.9 (d, *J* = 1.6 Hz); IR (film): $\tilde{\upsilon}$ = 2975, 2929, 1681, 1580, 1354, 1234, 1115, 972, 865, 813 cm^{-1}; MS (ESI): *m/z*: calcd for $C_9H_9FONa^+$: 175.1, found: 174.9.

1-(2,6-Dimethylphenyl)ethanone (9)

To a solution of 2-bromo-1,3-dimethylbenzene (**7**, 2 mL, 15.01 mmol) in anhydrous THF (20 mL) at -78 °C was added *n*BuLi (2.5 M in hexane, 9.6 mL, 24.02 mmol). The reaction was stirred 30 min at -78 °C. After addition of acetaldehyde (21.8 mL, 390.29 mmol), stirring was continued at -78 °C for 1 h. The reaction was quenched with water, and the organic layer was washed with brine, dried over $MgSO_4$, filtered, and concentrated under reduced pressure to afford the crude alcohol (**8**). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.07-7.03 (m, 1H), 7.00-6.98 (m, 2H), 5.40 (q, *J* = 6.7 Hz, 1H), 2.45 (s, 6H), 1.54 (d, *J* = 6.7 Hz, 3H), OH not observed; ^{13}C NMR (125 MHz, $CDCl_3$): δ = 140.5, 135.7, 129.4, 126.9, 67.7, 21.4, 20.6; IR (film): $\tilde{\upsilon}$ = 3302, 3019, 2925, 2856, 1583, 1468, 1366, 1298, 1191, 1071, 892 cm^{-1}; MS (ESI): *m/z*: calcd for $C_{10}H_{14}ONa^+$: 173.1, found: 173.0.

The crude intermediate (**8**) was dissolved in CH_2Cl_2 (20 mL), and PCC (3.56 g, 16.51 mmol) was added to the mixture. After 1 h at 25 °C, the solution was filtrated over celite, concentrated under reduced pressure, and purified by flash chromatography on silica gel (hexane/EtOAc 100:1) to afford a light yellow oil (1.24 g, 56% yield over two steps). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.15 (t, *J* = 7.6 Hz, 1H), 7.02 (d, *J* = 7.6 Hz, 2H), 2.47 (s, 3H), 2.25 (s, 6H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 208.4, 142.6, 132.2, 128.5, 127.8, 32.1, 19.1; IR (film): $\tilde{\upsilon}$ = 3066, 2954, 2923, 2854, 1697, 1462, 1421, 1350, 1254, 1055, 769 cm^{-1}; MS (ESI): *m/z*: calcd for $C_{10}H_{12}ONa^+$: 171.1, found: 171.0.

1-(3-Amino-2,6-dimethylphenyl)ethanone (11)

To a solution of 1-(2,6-dimethylphenyl)ethanone (**9**, 1.2 g, 8.10 mmol) in H_2SO_4 (10 mL) at 0 °C was added dropwise a solution of KNO_3 (819 mg, 8.10 mmol) in H_2SO_4 (10 mL) over 1 h. The reaction was then quenched by pouring into ice, and the aqueous layer was extracted with EtOAc. The organic solution was washed with brine, dried over $MgSO_4$, filtered, and concentrated under reduced pressure to afford the crude intermediate (**10**). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.81 (d, *J* = 8.3 Hz, 1H), 7.18 (d, *J* = 8.3 Hz, 1H), 2.51 (s, 3H), 2.42 (s, 3H), 2.31 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 205.9, 148.1, 144.7, 138.1, 128.9, 127.7, 124.5, 32.2, 19.4, 16.4; IR (film): $\tilde{\upsilon}$ = 3089, 2979, 2915, 2846, 1706, 1514, 1338, 1237, 1094, 841, 754 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{10}H_{11}NO_3Na^+$: 216.0631, found: 216.0631.

The intermediate (**10**) was dissolved in a 1 : 1 glacial acetic acid–ethanol solution (10 mL) and iron (2.26 g, 40.47 mmol) was added in portions. The resulting mixture was stirred at 90 °C for 2 h, filtered through celite, concentrated under reduced pressure, and neutralized with $NaHCO_3$. The aqueous layer was extracted with EtOAc, and the organic solution was dried over $MgSO_4$, filtered, and concentrated under reduced pressure. Purification by column chromatography on silica gel (hexane/EtOAc 5:1) afforded the desired compound as a light yellow oil (626 mg, 47% yield over two steps). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 6.85 (d, *J* = 8.1 Hz, 1H), 6.60 (d, *J* = 8.1 Hz, 1H), 3.57 (s, 2H), 2.45 (s, 3H), 2.14 (s, 3H), 2.03 (s, 3H); ^{13}C

NMR (125 MHz, $CDCl_3$): δ = 208.9, 143.2, 142.6, 128.6, 121.8, 116.4, 115.3, 32.4, 18.3, 13.8; IR (film): $\tilde{\upsilon}$ = 3455, 3369, 3235, 2919, 2862, 1691, 1624, 1482, 1350, 1302, 1242, 815 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{10}H_{13}NOH^+$: 164.1070, found: 164.1068.

1,1'-(5-Methyl-1H-indazole-1,4-diyl)diethanone (12)

To a mixture of 1-(3-amino-2,6-dimethylphenyl)ethanone (**11**, 285 mg, 1.74 mmol) and NaOAc (172 mg, 2.09 mmol) in $CHCl_3$ (3 mL) at 0 °C was added acetic anhydride (0.494 mL, 25.12 mmol) dropwise. After stirring for 30 min at 25 °C, the reaction was heated to 40 °C, and isoamyl nitrite (0.469 mL, 3.49 mmol) was added. The reaction was stirred at 60 °C for 7 h and quenched by the addition of a saturated solution of $NaHCO_3$. The resulting solution was extracted with $CHCl_3$, dried over $MgSO_4$, filtered, and concentrated under reduced pressure. Purification by column chromatography on silica gel (hexane/EtOAc 9:1) afforded the desired compound as a white solid (219 mg, 58% yield). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 8.42 (d, *J* = 8.5 Hz, 1H), 8.18 (d, *J* = 0.8 Hz, 1H), 7.42 (d, *J* = 8.5 Hz, 1H), 2.79 (s, 3H), 2.69 (s, 3H), 2.57 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 202.3, 171.2, 138.6, 137.8, 133.0, 132.4, 132.0, 124.2, 117.5, 32.1, 23.0, 20.3; IR (film): $\tilde{\upsilon}$ = 3107, 2979, 2927, 1719, 1686, 1412, 1380, 1356, 1160, 942, 818 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{12}H_{12}N_2O_2Na^+$: 239.0791, found: 239.0790.

General procedure for the α-bromination of acetophenones

A solution of the acetophenone (1 eq) in $CHCl_3$ (0.18 M) was added to a refluxing solution of copper (II) bromide (1.99 eq) in EtOAc (0.45 M). The mixture was then refluxed for 10 h. The solution was filtered through celite and concentrated under reduced pressure to afford a green solid. The solid was purified by flash chromatography on silica gel to afford the corresponding products in pure form. This method was used to obtain compounds **16**-**21**.

2-Bromo-1-(4-fluoro-2-methylphenyl)ethanone (16)

Yield: 64%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.73 (dd, *J* = 8.5 Hz, *J* = 5.7 Hz, 1H), 7.00-6.94 (m, 2H), 4.38 (s, 2H), 2.54 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 192.6, 164.5 (d, *J* = 254.8 Hz), 143.9 (d, *J* = 8.9 Hz), 131.8 (d, *J* = 9.7 Hz), 130.4 (d, *J* = 2.9 Hz), 119.3 (d, *J* = 21.2 Hz), 112.8 (d, *J* = 21.7 Hz), 33.2, 21.8 (d, *J* = 1.3 Hz); IR (film): $\tilde{\upsilon}$ = 2974, 2927, 1678, 1605, 1579, 1258, 1235, 1104, 989, 866, 818 cm^{-1}; MS (ESI): *m/z*: calcd for $C_9H_8BrFONa^+$: 253.0, found: 252.9.

2-Bromo-1-(2,6-dimethylphenyl)ethanone (17)

Yield: 21%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.23-7.19 (m, 1H), 7.06-7.04 (m, 2H), 4.28 (s, 2H), 2.26 (s, 6H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 200.1, 138.8, 133.4, 129.6, 127.9, 36.5, 19.4; IR (film): $\tilde{\upsilon}$ = 3066, 2925, 2854, 1716, 1695, 1462, 1381, 1261, 1185, 983, 771 cm^{-1}; MS (ESI): *m/z*: calcd for $C_{10}H_{11}BrONa^+$: 249.0, found: 248.9.

1-(1-Acetyl-5-methyl-1H-indazol-4-yl)-2-bromoethanone (18)

Yield: 37%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.47 (d, *J* = 8.6 Hz, 1H), 8.12 (s, 1H), 7.44 (d, *J* = 8.6 Hz, 1H), 4.43 (s, 2H), 2.80 (s, 3H), 2.57 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 195.8, 171.1, 138.0, 137.8, 132.8, 132.5, 129.2, 124.5, 118.2, 34.9, 23.0, 20.0; IR (film): $\tilde{\upsilon}$ = 2928, 2852, 1712, 1415, 1382, 1152, 946, 823, 677 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{12}H_{11}BrN_2O_2Na^+$: 316.9896, found: 316.9896.

2-Bromo-1-(2-chlorophenyl)ethanone (19)

Yield: 52%; ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.58-7.56 (m, 1H), 7.46-7.44 (m, 2H), 7.39-7.34 (m, 1H), 4.53 (s, 2H); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 194.0, 136.2, 132.7, 131.3, 130.6, 130.2, 127.1, 34.5; IR (film): $\tilde{\upsilon}$ = 3067, 2940, 1697, 1589, 1432, 1285, 1192, 1064, 988, 755 cm^{-1}; MS (ESI): *m/z*: calcd for $C_8H_6BrClONa^+$: 254.9, found: 254.9.

2-Bromo-1-(3-chlorophenyl)ethanone (20)

Yield: 31%; ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.95 (t, *J* = 1.9 Hz, 1H), 7.87-7.84 (m, 1H), 7.58 (ddd, *J* = 7.9 Hz, *J* = 1.9 Hz, *J* = 0.9 Hz, 1H), 7.44 (t, *J* = 7.9 Hz, 1H), 4.42 (s, 2H); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 190.1, 135.4, 135.2, 133.9, 130.2, 128.9, 127.0, 30.5; IR (film): $\tilde{\upsilon}$ = 3067, 3003, 2950, 1698, 1571, 1422, 1269, 1016, 781 cm^{-1}; MS (ESI): *m/z*: calcd for $C_8H_6BrClONa^+$: 254.9, found: 254.9.

2-Bromo-1-(2-(trifluoromethyl)phenyl)ethanone (21)

Yield: 55%; ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.75 (dd, *J* = 7.2 Hz, *J* = 1.5 Hz, 1H), 7.67-7.61 (m, 2H), 7.52 (dd, *J* = 7.2 Hz, *J* = 1.5 Hz, 1H), 4.40 (s, 2H); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 194.9, 136.3 (q, *J* = 2 Hz), 132.0, 131.1, 127.8, 127.4 (q, *J* = 32.4 Hz), 126.9 (q, *J* = 4.9 Hz), 123.2 (q, *J* = 273.6 Hz), 34.4 (q, *J* = 2.2 Hz); IR (film): $\tilde{\upsilon}$ = 2943, 1725, 1703, 1312, 1275, 1167, 1125, 1063, 1034, 766 cm^{-1}; MS (ESI): *m/z*: calcd for $C_9H_6BrF_3ONa^+$: 289.0, found: 288.9.

General procedure for the preparation of alkylated xanthines

3-Methyl-8-bromoxanthine (**23**, 1 eq) was dissolved in DMF (0.4 M), and *N,N*-diisopropylethylamine (1.5 eq) was added. After stirring for 5 min at 25 °C, α-bromoacetophenone (1 eq) was added. The reaction was stirred at 25 °C for 15 h. The mixture was then concentrated under reduced pressure, methanol was added and the formed precipitate was filtered off and washed with water to afford the corresponding product in pure form. This method was used to obtain compounds **24**-**28**.

8-Bromo-3-methyl-7-(2-oxo-2-(*o*-tolyl)ethyl)-1H-purine-2,6(3H,7H)-dione (24)

Yield: 58%; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.32 (s, 1H), 8.04 (dd, *J* = 7.6 Hz, *J* = 1.1 Hz, 1H), 7.56 (dt, *J* = 7.6 Hz, *J* = 1.1 Hz, 1H), 7.44 (t, *J* = 7.6 Hz, 1H), 7.40 (d, *J* = 7.6 Hz, 1H), 5.79 (s, 2H), 3.37 (s, 3H), 2.42 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 194.1, 153.9, 150.4, 148.9, 138.3, 133.7, 132.7, 131.9, 129.1,

129.0, 126.1, 108.8, 54.3, 28.5, 20.6; IR (film): $\tilde{\upsilon}$ = 3154, 3025, 2814, 1676, 1535, 1364, 1208, 768 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{15}H_{13}BrN_4O_3Na^+$: 399.0069, found: 399.0068.

8-Bromo-7-(2-(4-fluoro-2-methylphenyl)-2-oxoethyl)-3-methyl-1H-purine-2,6(3H,7H)-dione (25)

Yield: 56%; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.35 (s, 1H), 8.20-8.17 (m, 1H), 7.30-7.29 (m, 2H), 5.79 (s, 2H), 3.36 (s, 3H), 2.44 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 192.7, 164.0 (d, *J* = 252 Hz), 154.0, 150.5, 149.0, 142.9 (d, *J* = 9.2 Hz), 132.4 (d, *J* = 9.9 Hz), 130.3 (d, *J* = 2.7 Hz), 129.2, 118.8 (d, *J* = 21.5 Hz), 113.1 (d, *J* = 21.4 Hz), 108.8, 54.3, 28.6, 20.9; IR (film): $\tilde{\upsilon}$ = 3152, 3032, 2813, 1671, 1534, 1369, 1237, 1200, 990, 865 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{15}H_{12}BrFN_4O_3Na^+$: 416.9969, found: 416.9967.

8-Bromo-7-(2-(3-chlorophenyl)-2-oxoethyl)-3-methyl-1H-purine-2,6(3H,7H)-dione (26)

Yield: 75%; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.34 (s, 1H), 8.16 (t, *J* = 1.9 Hz, 1H), 8.07-8.05 (m, 1H), 7.85 (ddd, *J* = 7.9 Hz, *J* = 1.9 Hz, *J* = 0.9 Hz, 1H), 7.66 (t, *J* = 7.9 Hz, 1H), 5.96 (s, 2H), 3.36 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 190.7, 153.9, 150.5, 149.1, 135.4, 134.2, 134.0, 131.1, 129.1, 128.0, 126.9, 108.9, 53.0, 28.6; IR (film): $\tilde{\upsilon}$ = 3143, 3000, 2818, 1680, 1537, 1438, 1364, 1210, 876 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{14}H_{10}BrClN_4O_3Na^+$: 418.9517, found: 418.9515.

8-Bromo-7-(2-(2-chlorophenyl)-2-oxoethyl)-3-methyl-1H-purine-2,6(3H,7H)-dione (27)

Yield: 63%; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.38 (s, 1H), 7.96-7.94 (m, 1H), 7.68-7.65 (m, 2H), 7.61-7.56 (m, 1H), 5.80 (s, 2H), 3.36 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 193.0, 154.0, 150.5, 149.0, 134.4, 133.8, 131.0, 130.6, 130.1, 129.1, 127.7, 108.7, 54.9, 28.6; IR (film): $\tilde{\upsilon}$ = 3161, 3052, 2948, 2835, 1680, 1541, 1430, 1353, 1199, 838, 746 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{14}H_{10}BrClN_4O_3Na^+$: 418.9517, found: 418.9517.

8-Bromo-3-methyl-7-(2-oxo-2-(2-(trifluoromethyl)phenyl)ethyl)-1H-purine-2,6(3H,7H)-dione (28)

Yield: 54%; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.41 (s, 1H), 8.09 (d, *J* = 7.5 Hz, 1H), 7.96-7.92 (m, 2H), 7.87-7.84 (m, 1H), 5.78 (s, 2H), 3.36 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 194.8, 154.1, 150.5, 149.1, 135.2 (q, *J* = 2 Hz), 132.9, 132.3, 129.0, 128.5, 127.2 (q, *J* = 4.9 Hz), 126.2 (q, *J* = 32.2 Hz), 123.3 (q, *J* = 273.8 Hz), 108.7, 55.1, 28.6; IR (film): $\tilde{\upsilon}$ = 3163, 3046, 2940, 2807, 1708, 1674, 1359, 1318, 1213, 1165, 1111, 773 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{15}H_{10}BrF_3N_4O_3H^+$: 430.9961, found: 430.9963.

8-Bromo-7-(2-(2,6-dimethylphenyl)-2-oxoethyl)-3-methyl-1H-purine-2,6(3H,7H)-dione (29)

3-Methyl-8-bromoxanthine (**23**, 240 mg, 0.98 mmol) was dissolved in DMF (2.5 mL), and *N,N*-diisopropylethylamine (0.256 mL, 1.47 mmol) was added. After stirring for 5 min at 25 °C, 2-bromo-1-(2,6-dimethylphenyl)ethanone (**17**, 222 mg, 0.98 mmol) was added. The reaction was stirred at 110 °C for 2 h. The mixture was then concentrated under reduced pressure, methanol was added and the formed precipitate was filtered off and washed with water to afford the product in pure form (226 mg,

59% yield). ^{1}H NMR (400 MHz, DMSO-d_6): δ = 11.38 (s, 1H), 7.31 (t, *J* = 7.6 Hz, 1H), 7.15 (d, *J* = 7.6 Hz, 2H), 5.63 (s, 2H), 3.37 (s, 3H), 2.34 (s, 6H); ^{13}C NMR (125 MHz, DMSO-d_6, $CDCl_3$): δ = 201.0, 154.0, 150.4, 149.1, 137.5, 133.5, 129.7, 129.2, 127.8, 108.8, 56.5, 28.6, 18.8; IR (film): $\tilde{\upsilon}$ = 3150, 3019, 2956, 2829, 1668, 1363, 1199, 969, 780 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{16}H_{15}BrN_4O_3H^+$: 391.0400, found: 391.0405.

7-(2-(1-Acetyl-5-methyl-1H-indazol-4-yl)-2-oxoethyl)-8-bromo-3-methyl-1H-purine-2,6(3H,7H)-dione (30)

8-Bromo-3,9-dihydro-3-methyl-1H-purine-2,6-dione (**23**, 250 mg, 1.02 mmol) was added to a solution of KOH (57 mg, 1.02 mmol) in EtOH (3 mL). The resulting mixture was then heated to reflux for 2 h. EtOH was then removed under reduced pressure, and the resulting solid was washed with cold EtOH, and filtered off to afford a light yellow solid. A mixture of this solid (**18**, 144 mg, 0.51 mmol) was dissolved in DMF (1.3 mL), and 1-(1-acetyl-5-methyl-1H-indazol-4-yl)-2-bromoethanone (150 mg, 0.51 mmol) was added to the mixture. The reaction was stirred at 110 °C for 2 h. The mixture was then concentrated under reduced pressure, methanol was added and the formed precipitate was filtered off and washed with water to afford the product in pure form (184 mg, 39% yield over two steps). ^{1}H NMR (500 MHz, DMSO-d_6): δ = 11.48 (s, 1H), 8.64 (d, *J* = 0.7 Hz, 1H), 8.42 (d, *J* = 8.7 Hz, 1H), 7.65 (d, *J* = 8.7 Hz, 1H), 5.84 (s, 2H), 3.39 (s, 3H), 2.75 (s, 3H), 2.53 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 196.5, 170.8, 154.2, 150.5, 149.2, 137.8, 136.9, 133.0, 132.2, 129.3, 128.7, 123.5, 117.2, 108.8, 56.5, 28.6, 22.7, 19.0; IR (film): $\tilde{\upsilon}$ = 3165, 3049, 2948, 2820, 1722, 1702, 1675, 1371, 1207, 1154, 1105, 942, 816 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{18}H_{15}BrN_6O_4H^+$: 459.0411, found: 459.0407.

General procedure A for the cyclization of alkylated xanthines

A mixture of 3-alkyl-8-bromo-3,7-dihydro-7-(2-oxo-2-phenyl-ethyl)-1H-purine-2,6-dione (1.0 eq) and butylamine (0.1 M) was heated in a sealed tube at 175 °C for 30 min. The reaction was cooled to room temperature and evaporated to dryness. The resulting residue was triturated with a mixture acetone/water and the precipitate was filtered off. The crude intermediate was heated in a sealed tube at 175 °C in EtOH (0.1 M) in the presence of $AlCl_3$ (2 eq) for 2-15 h. The reaction was poured into water, extracted with CH_2Cl_2, and the organic layer was dried over $MgSO_4$, filtered, and concentrated under reduced pressure. Purification by column chromatography on silica gel (CH_2Cl_2/EtOAc 1:1) afforded the corresponding product in pure form. This method was used to obtain intermediates **31**-**36** and final products **40**-**45**.

General procedure B for the cyclization of alkylated xanthines

A mixture of 3-alkyl-8-bromo-3,7-dihydro-7-(2-oxo-2-phenyl-ethyl)-1H-purine-2,6-dione (1.0 eq) and the anisidine (4.0 eq) in EtOH (concentration, 0.1 M) was heated in a sealed tube at 175 °C for 12 h. The

reaction was cooled to room temperature, the formed solid was filtered off and washed with water to afford the corresponding product in pure form. This method was used to obtain compounds **47** and **48**.

8-(Butylamino)-7-(2-(butylimino)-2-(*o*-tolyl)ethyl)-3-methyl-1H-purine-2,6(3H,7H)-dione (31)

^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.53 (s, 1H), 7.31-7.25 (m, 4H), 6.90 (t, *J* = 5.6 Hz, 1H), 4.98 (d, *J* = 18.2 Hz, 1H), 4.85 (d, *J* = 18.2 Hz, 1H), 3.33 (s, 2H), 3.29 (s, 3H), 3.03-2.98 (m, 1H), 2.83-2.78 (m, 1H), 2.17 (s, 3H), 1.57-1.51 (m, 2H), 1.39-1.32 (m, 2H), 1.28-1.23 (m, 2H), 1.16-1.00 (m, 2H), 0.91 (t, *J* = 7.3 Hz, 3H), 0.69 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 163.7, 155.2, 153.1, 150.9, 150.0, 135.7, 134.4, 129.8, 128.6, 126.8, 125.7, 102.0, 51.0, 49.3, 41.9, 32.1, 31.4, 28.3, 19.5, 19.3, 18.6, 13.7, 13.6; IR (film): $\tilde{\upsilon}$ = 3329, 3141, 2954, 2929, 2869, 2806, 1671, 1611, 1573, 1438, 1300, 1225, 1113, 753 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{23}H_{32}N_6O_2H^+$: 425.2660, found: 425.2661.

3-Methyl-8-(pentylamino)-7-(2-(pentylimino)-2-(*o*-tolyl)ethyl)-1H-purine-2,6(3H,7H)-dione (32)

^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.53 (s, 1H), 7.31-7.25 (m, 4H), 6.90 (t, *J* = 5.6 Hz, 1H), 4.97 (d, *J* = 18.2 Hz, 1H), 4.86 (d, *J* = 18.2 Hz, 1H), 3.33 (s, 2H), 3.29 (s, 3H), 3.01-2.97 (m, 1H), 2.82-2.77 (m, 1H), 2.17 (s, 3H), 1.58-1.55 (m, 2H), 1.33-1.26 (m, 6H), 1.10-1.00 (m, 4H), 0.88 (t, *J* = 6.8 Hz, 3H), 0.75 (t, *J* = 6.9 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 163.6, 155.3, 153.1, 150.9, 150.0, 135.7, 134.4, 129.8, 128.6, 126.7, 125.7, 102.0, 51.1, 49.4, 42.2, 29.5, 28.9, 28.6, 28.4, 28.3, 21.9, 21.8, 18.6, 13.9, 13.8; IR (film): $\tilde{\upsilon}$ = 3340, 3144, 3015, 2955, 2924, 2855, 2802, 1657, 1614, 1579, 1437, 1219, 1098, 752 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{25}H_{36}N_6O_2H^+$: 453.2973, found: 453.2970.

8-(Hexylamino)-7-(2-(hexylimino)-2-(*o*-tolyl)ethyl)-3-methyl-1H-purine-2,6(3H,7H)-dione (33)

^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.52 (s, 1H), 7.31-7.25 (m, 4H), 6.90 (t, *J* = 5.6 Hz, 1H), 4.97 (d, *J* = 17.7 Hz, 1H), 4.86 (d, *J* = 17.7 Hz, 1H), 3.33 (s, 2H), 3.29 (s, 3H), 3.00-2.96 (m, 1H), 2.82-2.77 (m, 1H), 2.17 (s, 3H), 1.58-1.52 (m, 2H), 1.35-1.25 (m, 8H), 1.17-1.02 (m, 6H), 0.87 (t, *J* = 6.8 Hz, 3H), 0.78 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 163.6, 155.3, 153.1, 150.9, 150.0, 135.8, 134.4, 129.8, 128.6, 126.8, 125.6, 102.0, 51.1, 49.3, 42.2, 31.0, 30.9, 29.8, 29.2, 28.2, 26.0, 25.9, 22.1, 21.9, 18.6, 13.9, 13.8; IR (film): $\tilde{\upsilon}$ = 3338, 3149, 3015, 2953, 2924, 2854, 1659, 1614, 1550, 1486, 1439, 1301, 1218, 1099, 752 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{27}H_{40}N_6O_2H^+$: 481.3286, found: 481.3286.

8-(Butylamino)-7-(2-(butylimino)-2-(4-fluoro-2-methylphenyl)ethyl)-3-methyl-1H-purine-2,6(3H,7H)-dione (34)

^{1}H NMR (400 MHz, DMSO-d_6): δ = 10.50 (s, 1H), 7.37-7.32 (m, 1H), 7.15-7.13 (m, 2H), 6.87 (t, *J* = 5.7 Hz, 1H), 4.98 (d, *J* = 18.4 Hz, 1H), 4.83 (d, *J* = 18.4 Hz, 1H), 3.31 (s, 5H), 3.04-2.98 (m, 1H), 2.83-2.77 (m, 1H), 2.18 (s, 3H), 1.58-1.51 (m, 2H), 1.38-1.32 (m, 2H), 1.29-1.24 (m, 2H), 1.14-1.03 (m, 2H), 0.91 (t, *J* = 7.3 Hz, 3H), 0.70 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 162.9, 161.9 (d, *J* = 244.7 Hz), 155.2,

153.1, 150.9, 150.0, 137.8 (d, *J* = 8.2 Hz), 132.0 (d, *J* = 3.1 Hz), 129.0 (d, *J* = 8.8 Hz), 116.4 (d, *J* = 21.2 Hz), 112.6 (d, *J* = 21.0 Hz), 101.9, 51.0, 49.4, 41.9, 32.0, 31.4, 28.3, 19.5, 19.3, 18.6, 13.7, 13.6; IR (film): $\tilde{\upsilon}$ = 3330, 3141, 2956, 2930, 2870, 1672, 1656, 1612, 1573, 1439, 1300, 1226, 1116, 856, 754 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{23}H_{31}FN_6O_2H^+$: 443.2565, found: 443.2564.

8-(Butylamino)-7-(2-(butylimino)-2-(3-chlorophenyl)ethyl)-3-methyl-1H-purine-2,6(3H,7H)-dione (35)

^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.53 (s, 1H), 7.52-7.50 (m, 2H), 7.43 (s, 1H), 7.24-7.21 (m, 1H), 6.86 (t, *J* = 5.4 Hz, 1H), 4.97 (s, 2H), 3.31 (s, 2H), 3.28 (s, 3H), 3.08 (t, *J* = 6.6 Hz, 2H), 1.58-1.51 (m, 2H), 1.40-1.27 (m, 4H), 1.06-1.15 (m, 2H), 0.91 (t, *J* = 7.3 Hz, 3H), 0.71 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 162.0, 154.9, 153.2, 150.9, 150.0, 137.6, 133.3, 130.5, 128.7, 126.7, 125.6, 101.8, 51.2, 49.3, 41.9, 32.3, 31.4, 28.3, 19.5, 19.4, 13.7, 13.6; IR (film): $\tilde{\upsilon}$ = 3361, 3154, 2956, 2862, 1668, 1615, 1570, 1544, 1443, 1223, 749 cm^{-1}.

8-(Butylamino)-7-(2-(butylimino)-2-(2-chlorophenyl)ethyl)-3-methyl-1H-purine-2,6(3H,7H)-dione (36)

^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.52 (s, 1H), 7.53-7.39 (m, 4H), 6.81 (t, *J* = 5.4 Hz, 1H), 5.08 (d, *J* = 17.7 Hz, 1H), 4.92 (d, *J* = 17.7 Hz, 1H), 3.34 (s, 2H), 3.28 (s, 3H), 3.09-3.07 (m, 1H), 2.86-2.84 (m, 1H), 1.53 (quint, *J* = 7.3 Hz, 2H), 1.39-1.29 (m, 4H), 1.15-1.06 (m, 2H), 0.90 (t, *J* = 7.3 Hz, 3H), 0.71 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 160.8, 155.2, 153.1, 150.9, 150.0, 134.5, 130.6, 130.1, 129.1, 129.0, 127.3, 102.0, 51.5, 49.1, 41.9, 31.9, 31.3, 28.3, 19.5, 19.3, 13.7, 13.6; IR (film): $\tilde{\upsilon}$ = 3316, 3139, 2955, 2858, 2812, 1672, 1614, 1444, 1303, 1225, 1133, 875, 750 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{22}H_{29}ClN_6O_2H^+$: 445.2113, found: 445.2115.

8-Butyl-1-methyl-7-(*o*-tolyl)-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (40)

Yield: 32% over two steps; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.94 (s, 1H), 7.62 (s, 1H), 7.46-7.33 (m, 4H), 3.82 (t, *J* = 7.3 Hz, 2H), 3.40 (s, 3H), 2.24 (s, 3H), 1.54 (quint, *J* = 7.3 Hz, 2H), 1.06 (sext, *J* = 7.3 Hz, 2H), 0.68 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 153.3, 152.7, 151.0, 147.3, 138.2, 131.5, 130.4, 130.3, 129.8, 127.0, 125.9, 105.7, 98.9, 42.8, 30.1, 28.8, 19.6, 18.7, 13.0; IR (film): $\tilde{\upsilon}$ = 3165, 3062, 2955, 2872, 1698, 1670, 1505, 1456, 1146, 774, 721 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{19}H_{21}N_5O_2Na^+$: 374.1588, found: 374.1585.

1-Methyl-8-pentyl-7-(*o*-tolyl)-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (41)

Yield: 37% over two steps; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.94 (s, 1H), 7.62 (s, 1H), 7.47-7.33 (m, 4H), 3.82 (t, *J* = 7.1 Hz, 2H), 3.40 (s, 3H), 2.24 (s, 3H), 1.56 (quint, *J* = 7.1 Hz, 2H), 1.12-0.99 (m, 4H), 0.71 (t, *J* = 7.1 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 153.3, 152.7, 151.0, 147.3, 138.1, 131.5, 130.4, 130.3, 129.8, 127.1, 125.9, 105.7, 98.9, 43.0, 28.8, 27.6, 27.6, 21.2, 19.6, 13.5; IR (film): $\tilde{\upsilon}$ = 3156, 3013,

2930, 2868, 2820, 1674, 1514, 1455, 1297, 1153, 848, 725 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{20}H_{23}N_5O_2Na^+$: 388.1744, found: 388.1742.

8-Hexyl-1-methyl-7-(*o*-tolyl)-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (42)

Yield: 38% over two steps; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.94 (s, 1H), 7.62 (s, 1H), 7.46-7.33 (m, 4H), 3.82 (t, *J* = 7.1 Hz, 2H), 3.40 (s, 3H), 2.24 (s, 3H), 1.56 (quint, *J* = 7.1 Hz, 2H), 1.12-1.02 (m, 6H), 0.74 (t, *J* = 7.1 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 153.3, 152.7, 151.0, 147.3, 138.1, 131.5, 130.4, 130.3, 129.8, 127.1, 125.9, 105.7, 98.9, 43.0, 30.2, 28.8, 27.8, 25.1, 21.7, 19.6, 13.7; IR (film): $\tilde{\upsilon}$ = 3155, 3016, 2928, 2856, 2823, 1671, 1514, 1455, 1154, 851, 725 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{21}H_{25}N_5O_2Na^+$: 402.1901, found: 402.1905.

8-Butyl-7-(4-fluoro-2-methylphenyl)-1-methyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (43)

Yield: 43% over two steps; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.94 (s, 1H), 7.63 (s, 1H), 7.49-7.46 (m, 1H), 7.30 (d, *J* = 9.3 Hz, 1H), 7.22-7.19 (m, 1H), 3.81 (t, *J* = 7.1 Hz, 2H), 3.40 (s, 3H), 2.24 (s, 3H), 1.55 (quint, *J* = 7.1 Hz, 2H), 1.08 (sext, *J* = 7.1 Hz, 2H), 0.70 (t, *J* = 7.1 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 162.6 (d, *J* = 246.7 Hz), 153.3, 152.7, 151.0, 147.3, 141.5 (d, *J* = 8.5 Hz), 133.7 (d, *J* = 8.5 Hz), 129.2, 123.5 (d, *J* = 2.9 Hz), 116.9 (d, *J* = 21.5 Hz), 112.9 (d, *J* = 21.4 Hz), 106.0, 98.9, 42.8, 30.1, 28.8, 19.6, 18.8, 13.1; IR (film): $\tilde{\upsilon}$ = 3161, 3034, 2952, 2871, 2819, 1698, 1672, 1509, 1445, 1311, 1192, 1141, 861, 746 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{19}H_{20}FN_5O_2Na^+$: 392.1493, found:.392.1495.

8-Butyl-7-(3-chlorophenyl)-1-methyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (44)

Yield: 41% over two steps; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.98 (s, 1H), 7.84 (s, 1H), 7.72-7.71 (m, 1H), 7.58-7.56 (m, 3H), 4.12 (t, *J* = 7.3 Hz, 2H), 3.40 (s, 3H), 1.61 (quint, *J* = 7.3 Hz, 2H), 1.12 (sext, *J* = 7.3 Hz, 2H), 0.74 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 153.2, 152.9, 151.0, 148.0, 133.6, 130.8, 130.6, 130.1, 128.9, 128.5, 127.5, 106.0, 98.9, 43.5, 30.2, 28.9, 18.8, 13.1; IR (film): $\tilde{\upsilon}$ = 3153, 3027, 2969, 2929, 2859, 2826, 1686, 1507, 1442, 1226, 1138, 786 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{18}H_{18}ClN_5O_2H^+$: 372.1222, found: 372.1223.

8-Butyl-7-(2-chlorophenyl)-1-methyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (45)

Yield: 34% over two steps; ^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.97 (s, 1H), 7.72-7.69 (m, 2H), 7.64-7.59 (m, 2H), 7.55-7.52 (m, 1H), 3.87 (t, *J* = 7.3 Hz, 2H), 3.40 (s, 3H), 1.57 (quint, *J* = 7.3 Hz, 2H), 1.08 (sext, *J* = 7.3 Hz, 2H), 0.69 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 153.3, 152.9, 151.0, 147.3, 134.2, 133.6, 131.9, 129.8, 128.4, 127.6, 126.6, 106.5, 98.9, 43.3, 30.1, 28.9, 18.7, 13.0; IR (film): $\tilde{\upsilon}$ = 3176, 3064, 2963, 2875, 2804, 1698, 1669, 1504, 1455, 1148, 768 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{18}H_{18}ClN_5O_2H^+$: 372.1222, found: 372.1224.

7-(3-Chlorophenyl)-8-(2-methoxyphenyl)-1-methyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (47)

Yield: 45%; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 10.98 (s, 1H), 8.16 (s, 1H), 7.58 (dd, *J* = 7.7 Hz, *J* = 1.6 Hz, 1H), 7.54 (td, *J* = 7.7 Hz, *J* = 1.6 Hz, 1H), 7.38-7.30 (m, 3H), 7.22-7.19 (m, 2H), 7.14 (td, *J* = 7.6 Hz, *J* = 1.1 Hz, 1H), 3.58 (s, 3H), 3.28 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 154.7, 153.3, 152.9, 151.0, 148.1, 133.0, 131.5, 131.3, 130.3, 130.2, 129.8, 128.1, 126.9, 125.7, 122.3, 121.1, 113.0, 106.3, 99.2, 55.7, 28.9; IR (film): $\tilde{\upsilon}$ = 3169, 3022, 2826, 1708, 1670, 1505, 1434, 1161, 849, 740 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{21}H_{16}ClN_5O_3H^+$: 422.1014, found: 422.1012.

7-(2-Chlorophenyl)-8-(2-methoxyphenyl)-1-methyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (48)

Yield: 32%; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 11.00 (s, 1H), 7.89 (s, 1H), 7.52-7.46 (m, 2H), 7.43-7.36 (m, 3H), 7.31-7.27 (m, 1H), 7.08-7.02 (m, 2H), 3.57 (s, 3H), 3.28 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 154.3, 153.4, 152.9, 151.0, 147.3, 133.7, 133.0, 131.1, 131.1, 129.6, 129.5, 129.3, 126.8, 126.7, 121.7, 120.6, 112.5, 107.2, 99.1, 55.4, 28.8; IR (film): $\tilde{\upsilon}$ = 3150, 3044, 2821, 1665, 1509, 1490, 1422, 1280, 1159, 764 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{21}H_{16}ClN_5O_3H^+$: 422.1014, found: 422.1016.

8-(Butylamino)-3-methyl-7-(2-oxo-2-(2-(trifluoromethyl)phenyl)ethyl)-1H-purine-2,6(3H,7H)-dione (37)

A mixture of 8-bromo-3-methyl-7-(2-oxo-2-(2-(trifluoromethyl)phenyl)ethyl)-1H-purine-2,6(3H,7H)-dione (**28**, 490 mg, 1.13 mmol) and butyl amine (0.448 mL, 4.55 mmol) in EtOH (11 mL) was heated in a sealed tube at 175 °C for 1 h. The reaction was cooled to room temperature and the solid was filtered off to afford the crude product as a light yellow solid (345 mg, 74% yield). ^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.62 (s, 1H), 8.15 (d, *J* = 7.6 Hz, 1H), 7.94-7.89 (m, 2H), 7.85-7.82 (m, 1H), 7.03 (t, *J* = 5.4 Hz, 1H), 5.55 (s, 2H), 3.32 (s, 3H), 1.53 (quint, *J* = 7.3 Hz, 2H), 1.33 (sext, *J* = 7.3 Hz, 2H), 0.90 (t, *J* = 7.3 Hz, 3H), 3.31 (s, 2H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 195.2, 154.8, 153.2, 150.8, 150.2, 135.3, 132.6, 132.2, 129.3, 127.2 (q, *J* = 5.4 Hz), 126.4 (q, *J* = 32.2 Hz), 123.2 (q, *J* = 273.8 Hz), 101.6, 51.5, 41.9, 31.2, 28.3, 19.2, 13.6; IR (film): $\tilde{\upsilon}$ = 3379, 3153, 3035, 2939, 2875, 2804, 1671, 1618, 1573, 1547, 1317, 1223, 1170, 1130, 778, 750 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{19}H_{20}F_3N_5O_3H^+$: 424.1591, found: 424.1593.

8-(Butylamino)-7-(2-(2,6-dimethylphenyl)-2-oxoethyl)-3-methyl-1H-purine-2,6(3H,7H)-dione (38)

A mixture of 8-bromo-7-(2-(2,6-dimethylphenyl)-2-oxoethyl)-3-methyl-1H-purine-2,6(3H,7H)-dione (**29**, 88 mg, 0.22 mmol) and butyl amine (0.088 mL, 0.89 mmol) in EtOH (2 mL) was heated in a sealed tube at 175 °C for 2 h. The reaction was cooled to room temperature and the solid was filtered off to afford the crude product as a light yellow solid (68 mg, 81% yield). ^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.70 (s, 1H), 7.26 (t, *J* = 7.6 Hz, 1H), 7.11 (d, *J* = 7.6 Hz, 2H), 6.91 (t, *J* = 5.4 Hz, 1H), 5.35 (s, 2H), 3.34 (s, 5H), 2.27 (s, 6H), 1.52 (quint, *J* = 7.3 Hz, 2H), 1.34 (sext, *J* = 7.3 Hz, 2H), 0.90 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 202.7, 155.0, 153.2, 150.9, 150.2, 138.2, 133.5, 129.3, 127.7, 102.3, 53.3, 41.9, 31.1, 28.4,

19.3, 18.7, 13.6; IR (film): $\tilde{\upsilon}$ = 3342, 3152, 2987, 2934, 1673, 1614, 1545, 1439, 1223, 750 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{20}H_{25}N_5O_3H^+$: 384.2030, found: 384.2032.

8-(Butylamino)-3-methyl-7-(2-(5-methyl-1H-indazol-4-yl)-2-oxoethyl)-1H-purine-2,6(3H,7H)-dione (39)

A mixture of 7-(2-(1-acetyl-5-methyl-1H-indazol-4-yl)-2-oxoethyl)-8-bromo-3-methyl-1H-purine-2,6(3H,7H)-dione (**30**, 143 mg, 0.31 mmol) and butyl amine (0.123 mL, 1.24 mmol) in EtOH (3 mL) was heated in a sealed tube at 175 °C for 30 min. The reaction was cooled to room temperature and the solid was filtered off to afford the crude product as a light yellow solid (120 mg, 93% yield). ^{1}H NMR (500 MHz, DMSO-d_6): δ = 13.35 (s, 1H), 10.70 (s, 1H), 8.46 (s, 1H), 7.68 (d, *J* = 8.5 Hz, 1H), 7.34 (d, *J* = 8.5 Hz, 1H), 7.09 (t, *J* = 5.1 Hz, 1H), 5.57 (s, 2H), 3.34 (s, 5H), 2.47 (s, 3H), 1.55 (quint, *J* = 7.3 Hz, 2H), 1.36 (sext, *J* = 7.3 Hz, 2H), 0.91 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 197.2, 154.9, 153.2, 150.9, 150.3, 138.8, 132.6, 129.9, 129.4, 127.3, 120.4, 113.3, 101.7, 53.1, 41.9, 31.2, 28.4, 19.8, 19.3, 13.6; IR (film): $\tilde{\upsilon}$ = 3325, 3141, 3031, 2933, 2862, 2783, 1685, 1671, 1489, 1443, 1223, 1159, 1107, 921, 819, 748 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{20}H_{23}N_7O_3H^+$: 410.1935, found: 410.1937.

8-Butyl-1-methyl-7-(2-(trifluoromethyl)phenyl)-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (46)

8-(butylamino)-3-methyl-7-(2-oxo-2-(2-(trifluoromethyl)phenyl)ethyl)-1H-purine-2,6(3H,7H)-dione (**37**, 345 mg, 0.81 mmol) was heated in a sealed tube at 175 °C in EtOH (8 mL) in the presence of $AlCl_3$ (543 mg, 4.07 mmol) for 15 h. The reaction was poured in water, extracted with CH_2Cl_2, and the organic layer was dried over $MgSO_4$, filtered, and concentrated under reduced pressure. Purification by column chromatography on silica gel (CH_2Cl_2/EtOAc 1:1) afforded the desired compound as a white solid (50 mg, 15% yield). ^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.98 (s, 1H), 7.98 (d, *J* = 7.4 Hz, 1H), 7.87 (t, *J* = 7.4 Hz, 1H), 7.81 (t, *J* = 7.4 Hz, 1H), 7.77 (d, *J* = 7.4 Hz, 1H), 7.63 (s, 1H), 3.81 (t, *J* = 7.3 Hz, 2H), 3.40 (s, 3H), 1.56 (quint, *J* = 7.3 Hz, 2H), 1.13 (sext, *J* = 7.3 Hz, 2H), 0.72 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 153.3, 152.9, 151.0, 147.3, 134.2, 132.6, 130.8, 129.3 (q, *J* = 29.2 Hz), 127.0, 126.6 (q, *J* = 5.2 Hz), 125.4 (q, *J* = 1.7 Hz), 123.6 (q, *J* = 273.7 Hz), 106.6, 98.8, 43.2, 30.0, 28.9, 18.8, 13.1; IR (film): $\tilde{\upsilon}$ = 3169, 3034, 2959, 2873, 2820, 1701, 1671, 1509, 1443, 1309, 1179, 1136, 781 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{19}H_{18}F_3N_5O_2H^+$: 406.1485, found: 406.1483.

8-Butyl-7-(2,6-dimethylphenyl)-1-methyl-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (49)

8-(butylamino)-7-(2-(2,6-dimethylphenyl)-2-oxoethyl)-3-methyl-1H-purine-2,6(3H,7H)-dione (**38**, 175 mg, 0.46 mmol) was heated in a sealed tube at 175 °C in CH_2Cl_2 (3 mL) in the presence of $BF_3{\cdot}OEt_2$ (2 mL, 16.2 mmol) for 6 h. The reaction was poured in water, extracted with CH_2Cl_2, and the organic layer was dried over $MgSO_4$, filtered, and concentrated under reduced pressure. Purification by column chromatography on silica gel (EtOAc/Toluene 1:2) followed by recrystallization in acetone afforded the

desired compound as a white solid (43 mg, 26% yield). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.88 (s, 1H), 7.32 (t, *J* = 7.6 Hz, 1H), 7.26 (s, 1H), 7.18 (d, *J* = 7.6 Hz, 2H), 3.72 (t, *J* = 7.4 Hz, 2H), 3.60 (s, 3H), 2.15 (s, 6H), 1.65 (quint, *J* = 7.4 Hz, 2H), 1.23 (sext, *J* = 7.4 Hz, 2H), 0.82 (t, *J* = 7.4 Hz, 3H); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 153.6, 153.1, 151.1, 148.1, 139.4, 130.1, 129.4, 127.8, 126.5, 105.7, 99.8, 43.5, 30.9, 29.4, 20.4, 19.6, 13.4; IR (film): $\tilde{\upsilon}$ = 3152, 3044, 2958, 2871, 2820, 1675, 1607, 1512, 1457, 781 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{20}H_{23}N_5O_2H^+$: 366.1925, found: 366.1921.

8-Butyl-1-methyl-7-(5-methyl-1H-indazol-4-yl)-1H-imidazo[2,1-*f*]purine-2,4(3H,8H)-dione (50)

8-(butylamino)-3-methyl-7-(2-(5-methyl-1H-indazol-4-yl)-2-oxoethyl)-1H-purine-2,6(3H,7H)-dione (**39**, 40 mg, 0.09 mmol) was heated in a sealed tube at 175 °C in CH_2Cl_2 (4 mL) in the presence of $BF_3{\cdot}OEt_2$ (0.4 mL, 3.24 mmol) for 30 min. The reaction was poured in water, extracted with CH_2Cl_2, and the organic layer was dried over $MgSO_4$, filtered, and concentrated under reduced pressure. Purification by column chromatography on silica gel (gradient EtOAc:MeOH 99:1 to 98:2) afforded the desired compound as a white solid (1.1 mg, 3% yield). ^{1}H NMR (400 MHz, DMSO-d_6): δ = 13.21 (s, 1H), 10.93 (s, 1H), 7.83 (s, 1H), 7.69 (s, 1H), 7.63 (d, *J* = 8.5 Hz, 1H), 7.39 (d, *J* = 8.5 Hz, 1H), 3.77 (t, *J* = 7.3 Hz, 2H), 3.43 (s, 3H), 2.31 (s, 3H), 1.48 (quint, *J* = 7.3 Hz, 2H), 0.98 (sext, *J* = 7.3 Hz, 2H), 0.55 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 153.3, 152.8, 151.0, 147.6, 138.4, 132.2, 130.5, 128.9, 127.8, 124.3, 117.7, 111.8, 106.2, 99.0, 43.1, 30.0, 28.8, 18.8, 18.6, 12.9; IR (film): $\tilde{\upsilon}$ = 3356, 3194, 2956, 2925, 2857, 1687, 1607, 1513, 1456, 1054, 754 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{20}H_{21}N_7O_2Na^+$: 414.1649, found: 414.1646.

3.8.2 FRET based enzymatic assay

Compounds were tested in the Z'-LYTE™ Kinase Assay Kit–Tyr 1 Peptide (Invitrogen, USA) in a Corning 384 well microtiter plate. Fluorescence progress curves were measured upon excitation at 400 nm and emission at 445 and 520 nm. The assay contained a final concentration of EphB4 and ATP of 25 ng/µL and 125 µM (which is near its K_m), respectively, and was run at room temperature for 2 h. IC_{50} values (inhibitor concentration at which enzyme activity is reduced by 50%) are determined after carrying out assays at ten different concentrations between 20 µM and 10 pM.

3.8.3 [γ-33P]ATP based enzymatic assay

The enzymatic assays for the selectivity profile were performed at the National Centre for Protein Kinase Profiling at the University of Dundee. All assays (25.5 µL volume) were carried out robotically at room temperature and were linear with respect to time and enzyme concentration under the conditions used. Assays were performed for 30 min using Multidrop Micro reagent dispensers (Thermo Electron Corporation, Waltham, MA, USA) in a 96-well format. The concentration of magnesium acetate in the

assays was 10 mM and [γ-^{33}P]ATP (800 c.p.m./pmol) was used at 5, 20 or 50 μM in order to be at or below the K_m for ATP for each kinase. The assays were initiated with MgATP, stopped by the addition of 5 μL of 0.5 M orthophosphoric acid and spotted on to P81 filter plates using a unifilter harvester (PerkinElmer, Boston, MA, USA). The data is presented as mean percentage activity of duplicate assays at single concentration compared to DMSO controls. A similar protocol was used at Reaction Biology Corporation to measure the inhibitory activity of compounds **3**, **40**, **43** and **50** against Abl, Src, Lck and Yes1.

3.8.4 Cellular phosphorylation assays

The following experiments were performed at ProQinase GmbH. Mouse embryonal fibrablast cells, stably transfected to overexpress full-length human EphB4, were plated at 4×10^4 cells/well in DMEM supplemented with 10% FCS in 48-well culture dishes. Medium was replaced by DMEM without FCS before test compounds prediluted in 100% DMSO were added (final DMSO concentration of 1%). After incubation for 90 min at 37 °C, cells were stimulated for 2 h at 4 °C using murine ephrinB2-Fc at a final concentration of 2 μg/ml. Quantification of EphB4 phosphorylation was assessed in a 96-well plate via sandwich ELISA using a myc capture antibody and an antiphosphotyrosine detection antibody.

3.8.5 Cellular permeability assays

The following experiments were performed at Absorption Systems. Caco-2 monolayers were grown to confluence on collagen-coated, microporous, polycarbonate membranes in 12-well Costar Transwell® plates. The permeability assay buffer for the donor chambers was Hanks Balanced Salt Solution containing 10 mM HEPES and 15 mM glucose at a pH of 7.4. The buffer in the receiver chambers also contained 1% bovine serum albumin. Cells were dosed on the apical side (A-to-B) or basolateral side (B-to-A) and incubated at 37 °C with 5% CO_2 in a humidified incubator. After 2 h, aliquots were taken from the donor and receiver chambers. Each determination was performed in duplicate. The flux of lucifer yellow was also measured for each monolayer after being subjected to the test compounds to ensure that no damage was inflicted to the cell monolayers during the flux period. All samples were assayed by LC-MS/MS using electrospray ionization. The apparent permeability, P_{app}, and percent recovery were calculated as follows:

$$P_{app} = (dC_r/dt) \times V_r/(A \times C_A) \quad (1)$$

$$\text{Percent Recovery} = 100 \times ((V_r \times C_r^{final}) + (V_d \times C_d^{final}))/(V_d \times C_0) \quad (2)$$

where, dC_r/dt is the slope of the cumulative concentration in the receiver compartment versus time in $\mu M.s^{-1}$, V_r is the volume of the receiver compartment in cm^3, V_d is the volume of the donor compartment in cm^3, A is the area of the cell monolayer (1.13 cm^2 for 12-well Transwell®), C_0 is the

nominal concentration of the dosing solution in µM, C_A is the average of the nominal concentration of the dosing solution and the measured donor concentration at 120 min in µM, C_r^{final} is the cumulative receiver concentration in µM at the end of the incubation period, C_d^{final} is the concentration of the donor in µM at the end of the incubation period.

3.8.6 X-ray crystallography

Protein expression and purification

A clone of the EphA3 kinase domain (residues: 606-947) was obtained from Prof. Sirano Dhe-Paganon's group[16] and expressed in *Escherichia coli* strain BL21 (DE3). Cells expressing EphA3 were induced with a 1 mM solution of isopropyl-beta-D-thiogalactopyranoside (IPTG) for 12 h at 15 °C. Cell pellets were resuspended in buffer A (50 mM Tris, pH 8.0, and 100 mM NaCl, supplemented with protease inhibitors) and lysed by sonication. After centrifugation at 15,000 rpm for 1 h, the soluble fraction of EphA3 was purified using HisTrap FF crude and HiTrap Q HP columns (GE Healthcare), followed by gel filtration chromatography (Superdex75; GE Healthcare). The appropriate fractions were combined and concentrated to ~10 mg/mL using Amicon filter devices (10 kDa as cutoff) in a storage solution (100 mM sodium chloride and 10 mM Tris-HCl pH 8.0, 5% glycerol). The resulting solution was aliquoted and stored at -80 °C for further usage.

Crystallization, Data Collection, and Structure Determination

Crystals of the EphA3 kinase domain were grown at 20 °C using the hanging drop vapor diffusion method. Equal volumes of protein and reservoir solutions (0.1 M sodium cacodylate pH 6.5, 0.15 M ammonium sulfate, 22.5% PEG 3350) were mixed and crystals appeared after 1 to 2 days. A 5 mM solution of inhibitor (in 100% DMSO) was added into the hanging drop to reach a final DMSO concentration of 10% (v/v). The crystals were soaked for 1 to 24 h and flash-frozen in liquid nitrogen without extra cryoprotectant.

Refinement and validation

Data sets were collected on a MarCCD detector and indexed, integrated and scaled with the XDS[17] and CCP4 programs[18]. The structures were solved by molecular replacement with PHASER[19] using the apo EphA3 kinase domain structure (PDB entry 2GSF) as a search model and refined with PHENIX.[20]

3.9 Additional experimental data

Table 5. Selectivity of compound **40** measured at the University of Dundee.

Kinase	Percentage of kinase activity compared to a 100% DMSO control
MKK1	56
MKK2	86
MKK6	105
ERK1	96
ERK2	105
JNK1	96
JNK2	95
JNK3	87
p38a MAPK	107
p38b MAPK	98
p38g MAPK	108
p38d MAPK	102
ERK8	94
RSK1	98
RSK2	84
PDK1	104
PKBa	103
PKBb	74
SGK1	134
S6K1	114
PKA	83
ROCK 2	101
PRK2	100
PKCa	105
PKCγ	99
PKCz	101
PKD1	80
STK33	84

Kinase	Percentage of kinase activity compared to a 100% DMSO control
MSK1	85
MNK1	105
MNK2	103
MAPKAP-K2	99
MAPKAP-K3	70
PRAK	99
CAMKKb	99
CAMK1	84
SmMLCK	104
PHK	93
DAPK1	115
CHK1	108
CHK2	107
GSK3b	123
CDK2-Cyclin A	114
PLK1	102
Aurora A	77
Aurora B	95
TLK1	116
LKB1	91
AMPK	100
MARK1	101
MARK2	111
MARK3	84
MARK4	97
BRSK1	92
BRSK2	83
MELK	101
NUAK1	109
TSSK1	98

Kinase	Percentage of kinase activity compared to a 100% DMSO control
CK1	97
CK2	91
TTBK1	97
DYRK1A	96
DYRK2	93
DYRK3	101
NEK2a	108
NEK6	90
IKKb	80
IKKe	93
TBK1	101
PIM1	98
PIM2	96
PIM3	92
SRPK1	91
EF2K	97
EIF2AK3	96
HIPK1	120
HIPK2	87
HIPK3	87
CLK2	118
PAK2	100
PAK4	106
PAK5	104
PAK6	96
MST2	111
MST4	103
GCK	93
MINK1	99
MEKK1	112

Kinase	Percentage of kinase activity compared to a 100% DMSO control
MLK1	106
MLK3	101
TESK1	84
TAO1	107
ASK1	111
TAK1	107
IRAK1	118
IRAK4	97
RIPK2	12
OSR1	104
TTK	104
MPSK1	83
Src	2
Lck	5
CSK	57
YES1	14
ABL	43
BTK	58
JAK2	108
SYK	137
ZAP70	103
TIE2	101
BRK	79
EPH-A2	6
EPH-A4	5
EPH-B1	10
EPH-B2	8
EPH-B3	13
EPH-B4	18
FGF-R1	88

Kinase	Percentage of kinase activity compared to a 100% DMSO control
HER4	91
IGF-1R	112
IR	99
IRR	111
TrkA	90
VEG-FR	87

3.10 References

(1) Kolb, P.; Kipouros, C. B.; Huang, D.; Caflisch, A. Structure-based tailoring of compound libraries for high-throughput screening: discovery of novel EphB4 kinase inhibitors. *Proteins* **2008**, *73*, 11-18.

(2) Lafleur, K.; Huang, D.; Zhou, T.; Caflisch, A.; Nevado, C. Structure-based optimization of potent and selective inhibitors of the tyrosine kinase erythropoietin producing human hepatocellular carcinoma receptor B4 (EphB4). *J. Med. Chem.* **2009**, *52*, 6433-6446.

(3) Bohm, H. J.; Banner, D.; Bendels, S.; Kansy, M.; Kuhn, B.; Muller, K.; Obst-Sander, U.; Stahl, M. Fluorine in medicinal chemistry. *Chembiochem* **2004**, *5*, 637-643.

(4) Auffinger, P.; Hays, F. A.; Westhof, E.; Ho, P. S. Halogen bonds in biological molecules. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16789-16794.

(5) Bamborough, P.; Angell, R. M.; Bhamra, I.; Brown, D.; Bull, J.; Christopher, J. A.; Cooper, A. W.; Fazal, L. H.; Giordano, I.; Hind, L.; Patel, V. K.; Ranshaw, L. E.; Sims, M. J.; Skone, P. A.; Smith, K. J.; Vickerstaff, E.; Washington, M. N-4-Pyrimidinyl-1H-indazol-4-amine inhibitors of Lck: indazoles as phenol isosteres with improved pharmacokinetics. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 4363-4368.

(6) Ishikawa, M.; Hashimoto, Y. Improvement in aqueous solubility in small molecule drug discovery programs by disruption of molecular planarity and symmetry. *J. Med. Chem.* **2011**, *54*, 1539-1554.

(7) Maier, J. A.; Brugel, T. A.; Sabat, M.; Golebiowski, A.; Laufersweiler, M. J.; VanRens, J. C.; Hopkins, C. R.; De, B.; Hsieh, L. C.; Brown, K. K.; Easwaran, V.; Janusz, M. J. Development of N-4,6-pyrimidine-N-alkyl-N'-phenyl ureas as orally active inhibitors of lymphocyte specific tyrosine kinase. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3646-3650.

(8) Sabat, M.; VanRens, J. C.; Laufersweiler, M. J.; Brugel, T. A.; Maier, J.; Golebiowski, A.; De, B.; Easwaran, V.; Hsieh, L. C.; Walter, R. L.; Mekel, M. J.; Evdokimov, A.; Janusz, M. J. The development of 2-

benzimidazole substituted pyrimidine based inhibitors of lymphocyte specific kinase (Lck). *Bioorg. Med. Chem. Lett.* **2006**, *16*, 5973-5977.

(9) Zagorska, A.; Jurczyk, S.; Pawlowski, M.; Dybala, M.; Nowak, G.; Tatarczynska, E.; Nikiforuk, A.; Chojnacka-Wojcik, E. Synthesis and preliminary pharmacological evaluation of imidazo[2,1-f]purine-2,4-dione derivatives. *Eur. J. Med. Chem.* **2009**, *44*, 4288-4296.

(10) Baraldi, P. G.; Preti, D.; Tabrizi, M. A.; Fruttarolo, F.; Romagnoli, R.; Zaid, N. A.; Moorman, A. R.; Merighi, S.; Varani, K.; Borea, P. A. New pyrrolo[2,1-f]purine-2,4-dione and imidazo[2,1-f]purine-2,4-dione derivatives as potent and selective human A3 adenosine receptor antagonists. *J. Med. Chem.* **2005**, *48*, 4697-4701.

(11) Baraldi, P. G.; Preti, D.; Tabrizi, M. A.; Romagnoli, R.; Saponaro, G.; Baraldi, S.; Botta, M.; Bernardini, C.; Tafi, A.; Tuccinardi, T.; Martinelli, A.; Varani, K.; Borea, P. A. Structure-activity relationship studies of a new series of imidazo[2,1-f]purinones as potent and selective A(3) adenosine receptor antagonists. *Bioorg. Med. Chem.* **2008**, *16*, 10281-10294.

(12) Messaoudi, S.; Treguier, B.; Hamze, A.; Provot, O.; Peyrat, J. F.; De Losada, J. R.; Liu, J. M.; Bignon, J.; Wdzieczak-Bakala, J.; Thoret, S.; Dubois, J.; Brion, J. D.; Alami, M. Isocombretastatins a versus combretastatins a: the forgotten isoCA-4 isomer as a highly promising cytotoxic and antitubulin agent. *J. Med. Chem.* **2009**, *52*, 4538-4542.

(13) Alcalde, E.; Mesquida, N.; Frigola, J.; Lopez-Perez, S.; Merce, R. Indene-based scaffolds. Design and synthesis of novel serotonin 5-HT6 receptor ligands. *Org. Biomol. Chem.* **2008**, *6*, 3795-3810.

(14) Takami, A.; Iwakubo, M.; Okada, Y.; Kawata, T.; Odai, H.; Takahashi, N.; Shindo, K.; Kimura, K.; Tagami, Y.; Miyake, M.; Fukushima, K.; Inagaki, M.; Amano, M.; Kaibuchi, K.; Iijima, H. Design and synthesis of Rho kinase inhibitors (I). *Bioorg. Med. Chem.* **2004**, *12*, 2115-2137.

(15) Bakke, B. A.; McIntosh, M. C.; Turnbull, K. D. Improved alkylation and product stability in phosphotriester formation through quinone methide reactions with dialkyl phosphates. *J. Org. Chem.* **2005**, *70*, 4338-4345.

(16) Choi, Y.; Syeda, F.; Walker, J. R.; Finerty, P. J., Jr.; Cuerrier, D.; Wojciechowski, A.; Liu, Q.; Dhe-Paganon, S.; Gray, N. S. Discovery and structural analysis of Eph receptor tyrosine kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 4467-4470.

(17) Kabsch, W. Automatic Processing of Rotation Diffraction Data from Crystals of Initially Unknown Symmetry and Cell Constants. *J. Appl. Crystallogr.* **1993**, *26*, 795-800.

(18) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **1994**, *50*, 760-763.

(19) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40*, 658-674.

(20) Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L. W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **2002**, *58*, 1948-1954.

Chapter 4: Lead optimization and fluorescence properties of novel EphB4 inhibitors

4 Lead optimization and fluorescence properties of novel EphB4 inhibitors

4.1 Introduction

Recently, novel type I EphB4 inhibitors were discovered by *in silico* flexible ligand docking[1] (as opposed to fragment-based docking) starting from a library of about 2 million compounds. A focused library was first generated by discarding molecules unable to interact with Met696 or to occupy the deep ATP back pocket. In addition, particular attention was given to the process of ligand binding involving the displacement of water molecules within the binding site and the formation of new hydrogen bonds with the receptor. In fact, binding of the ligand to the protein is unlikely if the newly formed hydrogen bonds are weaker, and an estimation of the hydrogen bonding penalty may allow the removal of unrealistic poses. Hence, a new approach was developed to estimate hydrogen bonding penalty upon ligand binding, and introduction of this parameter within the scoring function provided a better estimation of the binding energy. Among the 22,517 molecules having a calculated binding energy lower than -6 kcal/mol, only 1,381 made a hydrogen bond with Glu694. Further visual inspection and biochemical characterization lead to the discovery of compounds **1** and **2**, with IC_{50} values around 300 nM. The major features of their binding mode obtained by automatic docking are indicated in Figure 1, and include hydrogen bonding interactions with Met696, Glu694 and Thr693. To our delight, compounds **1** and **2** also exhibited interesting fluorescent properties, making them attractive candidates for further lead optimization.

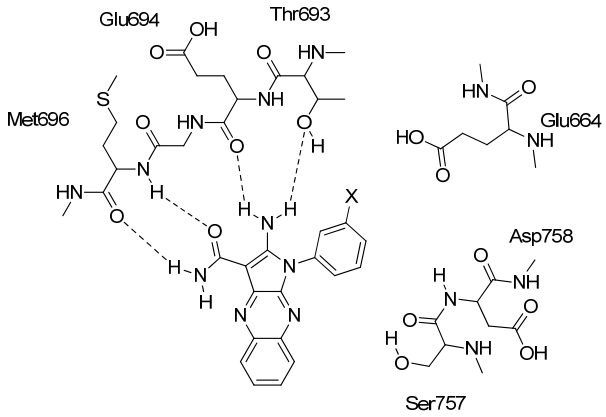


Figure 1. Binding mode of compound **1** (X = OMe) and **2** (X = F).

The development of fluorescent probes has been recognized as an efficient strategy to understand protein distribution and chemical environment, and has culminated in the discovery of green fluorescent

proteins (GFPs).[2,3] This powerful method consists in the fusion of a gene encoding a GFP and one encoding the target protein, resulting in the expression of a fluorescent fusion protein.

In an attempt to further understand protein function, small fluorescent molecules such as GSTO1[4] or dihydroflate reductase inhibitors[5] have recently been developed. In addition, chemical probes have been shown to be extremely useful for the characterization of drug repartition, uptake and mode of action.[6,7] Recently, the development of fluorescent COX2 inhibitors allowed *in vitro* and *in vivo* monitoring of the drug distribution.[8] Furthermore, the behavior of micelle drug carriers has been very well characterized using an internal fluorescent probe, as drug delivery was associated with significant fluorescence emission.[9] A better understanding of drug distribution is crucial for the development and optimization of new biological active compounds. For example, the synthesis of fluorescent cis-platine derivatives has been described, giving a detailed overview of the drug behavior within the cell.[10]

The synthesis of fluorescent probes can be easily achieved by attaching fluorophores to the molecule. However, this strategy has major drawbacks, such as toxicity or weaker stability,[10] and the discovery of emissive drugs still represents a very attractive strategy for the monitoring of biological processes.

Here, our aim was to optimize the binding affinity of moderate EphB4 inhibitors recently discovered by *in silico* high throughput screening,[1] and to use their fluorescence properties for cellular imaging.* Furthermore, the validation of their binding mode was investigated using X-ray crystallography.

4.2 Confirmation of the binding mode by X ray crystallography

In order to confirm the binding mode of compounds **1** and **2** obtained by automatic docking, X-ray crystallography studies were investigated (performed by Dr. Jing Dong, Table 1). As EphA3 could be produced in Escherichia coli with high yield, we decided to focus on the crystallization of this protein kinase whose structure resolution seemed also higher than EphB4. The main parameters of the X-ray structures obtained for compounds **1** and **2** using the hanging drop vapor diffusion method have been summarized in Table 1 (see also section 4.10.6).

* Due to the preparation of a patent application, most of the structures could not be disclosed.

Table 1. X-ray data collection.

	1	**2**
Space group	P 1 21 1	P 1 21 1
Unit cell		
a (Å)	53.91	53.51
b (Å)	38.16	38.17
c (Å)	75.53	75.66
Resolution range (Å)	38.16-1.6	47.5-1.8
Unique reflections	39425 (5697)	28152 (3794)
<I/σ(I)>	13.9 (3.7)	15.0 (3.7)
R merge	0.052 (0.317)	0.051 (0.293)
Completeness (%)	98.2 (97.6)	98.3 (91.0)
Multiplicity	3.5 (3.5)	3.6 (3.1)
Refinement		
Resolution range (Å)	33.64-1.60	29.76-1.79
R factor / R free	17.70 / 20.18	17.72 / 20.58
Mean B factors ($Å^2$)	21.20	25.50
RMS bonds (Å)	0.0064	0.0073
RMS angles (°)	1.286	1.340

4.3 Comparison of the docking pose with the binding mode obtained by X-ray crystallography

The most important features of the binding mode obtained by automatic docking were confirmed by X-ray crystallography. As predicted, the phenyl ring of compounds **1** and **2** is located into the hydrophobic pocket of the ATP binding site. The amino group of the molecule is involved in two hydrogen bonds with the backbone carbonyl of Glu694 and the side chain of Thr693. In addition, the hydrogen bonding interactions between the amide moiety and the backbone polar groups of Met696 are confirmed. The remarkable similarity between the docking pose and the experimental binding mode is highlighted in Figure 2.

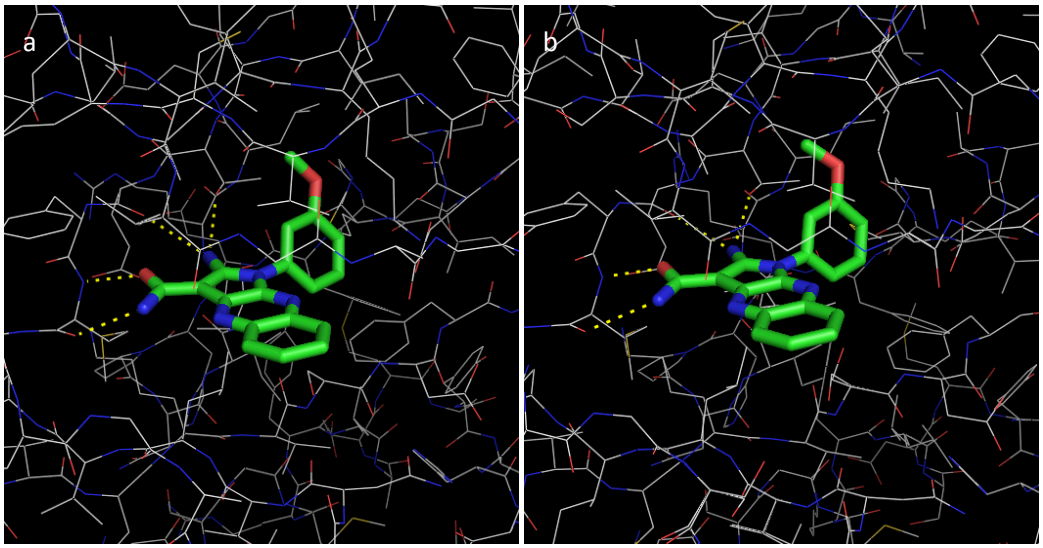


Figure 2. (a) Binding mode of compound **1** obtained by automatic docking on EphB4. (b) Binding mode of compound **1** obtained by X-ray crystallography on EphA3.

4.4 Lead optimization

According to the binding mode obtained by automatic docking and X-ray crystallography, the phenyl ring of compounds **1** and **2** is buried into the hydrophobic pocket of the ATP binding site (Figure 1). Thus, in order to increase van der Waals interactions with the side chains of Val629, Ala645, and Thr693, we decided to start with the addition of small alkyl groups. These simple substitutions appeared to be the most suitable from a synthetic point of view.

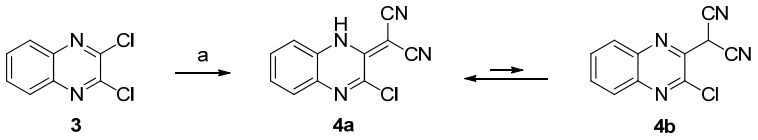
A few other residues within the binding pocket already proved to act as attractive hydrogen bonding partners. In particular, the formation of additional hydrogen bonds between a moderate inhibitor and the side chain of Glu664 has already resulted in single digit nanomolar binding affinity.[11-15] Therefore, we investigated the synthesis of compounds bearing hydrogen bond donors on the phenyl ring.

Finally, the synthesis of bis-orthosubstituted derivatives was investigated, as a rigidification of biologically active molecules already proved to be beneficial in terms of selectivity and solubility properties.[11] Increase in steric hindrance limits the rotation of the phenyl ring, and restricts the number of accessible conformations adopted by the molecule, thus limiting interactions with secondary targets. Disruption of molecular planarity has already been shown to decrease crystal packing and improve aqueous solubility.[16]

4.5 Synthesis

The synthesis of 1H-pyrrolo[2,3-b]quinoxaline derivatives started with commercially available 2,3-dichloroquinoxaline **3**, which reacted with malononitrile in the presence of sodium hydride to give intermediate **4a** (Scheme 1).[17,18] As previously reported, **4a** also exists as its tautomeric form **4b**, but the presence of a weak signal at 44 ppm in the ^{13}C-NMR spectrum suggests that it mainly occurs as its methylidene form.[18]

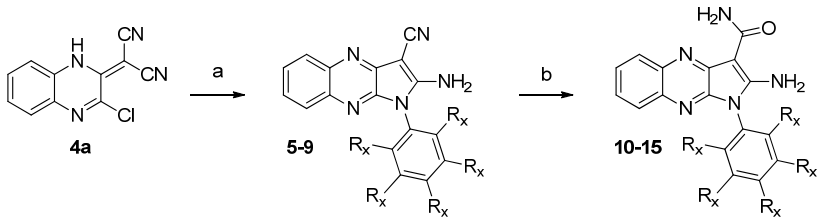
Scheme 1[a]



[a]Reagents and conditions: (a) (i) Malononitrile, NaH, dimethoxyethane, 25 °C, 30 min. ; (ii) (**3**), 25 °C, 3 h then reflux, 1 h.

Substitution of the chlorine at position 3 by a primary amine followed by intramolecular cyclization afforded intermediates **5** to **9** (Scheme 2).[19] The presence of a cyano group was confirmed by the presence of a characteristic IR band at around 2200 cm^{-1} and a ^{13}C-NMR signal at 115-116 ppm. As previously observed, the amino group appeared as a broad signal at 8 to 8.5 ppm.[19] Subsequent hydrolysis of the cyano group in the presence of concentrated sulfuric acid occurred at room temperature to afford final amides **10** to **15**.

Scheme 2[a]



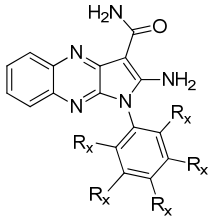
[a]Reagents and conditions: (a) Primary aromatic amine, EtOH, Toluene, 130-160 °C, 3 h. (b) H_2SO_4, 25 °C, 30 min.

4.6 Enzymatic and cellular phosphorylation assay

The inhibition activity of compounds **10**-**15** on EphB4 was assessed by FRET based enzymatic assay (Table 2). While the addition of a small alkyl group seemed quite promising (compound **10**, 66% inhibition at 1 µM), the addition of a hydrogen bond acceptor (compound **11**) significantly reduced the

binding affinity. In addition, bis-orthosubstituted derivatives (compounds **13** and **14**) showed micromolar activity. However, the introduction of hydrogen bond donors (compounds **12** and **15**) lead to a significant increase in binding affinity, with compound **15** showing double digit nanomolar inhibition. Furthermore, the inhibition activity of compounds **10**-**12** and **15** was confirmed by a competition binding assay on EphB4 and EphA3 (Ambit Bioscience Corporation). Compound **15** showed the highest level of activity, with a binding constant of 0.97-1.1 nM on EphA3.

Table 2. Results on FRET based enzymatic assays and competition binding assays.



Compound	R_x	R_x	R_x	R_x	R_x	FRET based assay (@ 1 µM)[a]	Competition binding assay (@ 1 µM)[b]	
							EphB4	EphA3
10	Alkyl	H	H	H	H	66%	48%	36%
11	Alkyl	H	H	HBA[c]	H	26%	79%	55%
12	Alkyl	H	H	HBD[d]	H	67%	40%	5.8%
13	Alkyl	H	H	H	Alkyl	32%	n.d.	n.d.
14	Alkyl	H	H	H	Halogen	41%	n.d.	n.d.
15	Alkyl	H	H	OH	H	100% (11 nM)	0.1%	0% (0.97-1.1 nM)

[a]Percentage of inhibition activity.
[b]Percentage of kinase activity compared to a 100% DMSO control.
[c]HBA = hydrogen bond acceptor.
[d]HBD = hydrogen bond donor.

In addition, the activity of compounds **10**-**15** was measured by cellular phosphorylation assays (Table 3, performed at Proqinase) on murine embryonal fibroblast cells (MEF) transfected with myc-tagged human EphB4. Quantification of EphB4 phosphorylation was assessed via sandwich ELISA after 90 min of incubation with the corresponding inhibitor. While bis-substituted compounds **13** and **14** showed micromolar activity, compounds **10**-**12** and **15** had a cellular IC_{50} below 800 nM. High cell membrane permeability was observed for compound **15**, with a cellular IC_{50} below 6 nM.

Table 3. Cellular phosphorylation assay.

Compound	Cell IC_{50} (nM)
10	230
11	720
12	160
13	4400
14	2800
15	< 6

4.7 Measurement of fluorescence emission properties

Fluorescence emission properties of compounds **10**-**15** were evaluated (Table 4). Each compound had similar absorption (λ_{abs} = 370 nm) and emission properties (λ_{em} = 430 nm), making them perfectly suitable for cellular imaging. In addition, their quantum yields were measured in acetonitrile as described,[20] providing values comprised between 0.20 and 0.25.

Table 4. Quantum yield measurements.

Compound	φ
10	0.20
11	0.22
12	0.23
13	0.21
14	0.22
15	0.25

4.8 Anti-proliferative effect of compound 15 on human breast cancer cells and fluorescence microscopy

The anti-proliferative activity of a previously reported inhibitor of EphB4, dasatinib, has been assessed on MDA-MB-231 human breast cancer cells (IC_{50} = 12 nM after 72 hours of incubation).[21] Interestingly, gene expression profiling of MDA-MB-231 cells provided evidence for up-regulation of EphB2, EphA2, EphA5, EphA7 and EphA10 receptor tyrosine kinases.[22] Thus, we decided to measure the

inhibitory effect of compound **15** on cell growth using a resazurin colorimetric assay. Significant anti-proliferative effects were observed at micromolar concentrations (IC_{50} = 8-11 µM, Figure 3) after incubation for 24 hours.

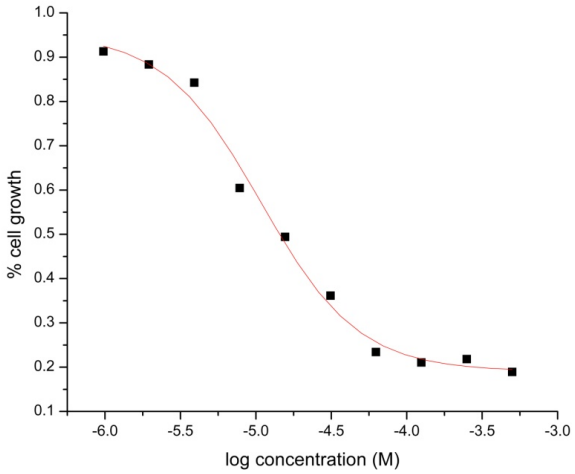


Figure 3. Cellular proliferation assay on MDA-MB-231 (compound **15**).

After 2 hours of incubation, significant changes in cell morphology were observed, such as cell shrinkage and disruption of cell-cell adhesion (compare Figure 4a and 4b). Similar results have been obtained after incubation with the CDK inhibitor roscovitine,[23] and have been shown to be associated with apoptosis.

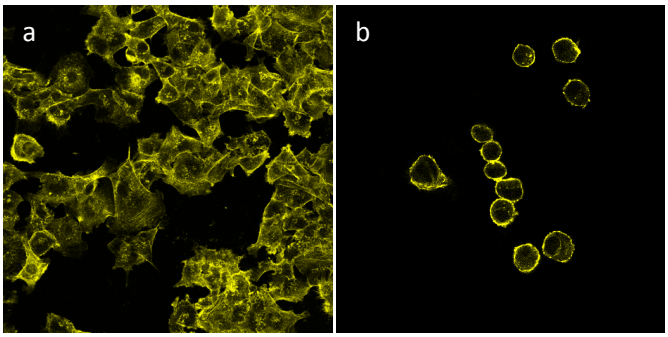


Figure 4. Fluorescence images of MDA-MB-231 cells stained with tetramethylrhodamine isothiocyanate (TRITC)-phalloidin. (a) Control cells. (b) After 2 h of incubation with compound **15** (25 µM).

In addition, fluorescence emission of compound **15** was detected within the cytoplasm after 2 hours of incubation (Figure 5), which confirms a remarkable cell membrane permeability.

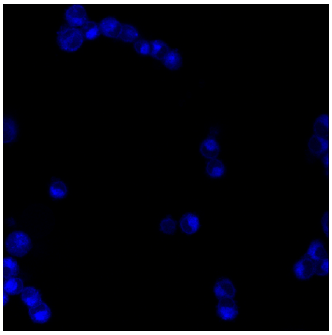


Figure 5. Fluorescence images of MDA-MB-231 cells after 2 h of incubation with compound **15** (25 µM, no staining).

4.9 Conclusions

Ligand binding is a complex event inducing the displacement of water molecules and the formation of new hydrogen bonds with the receptor. In fact, molecules involving weaker hydrogen bonding interactions are likely to be inactive, and the recent evaluation of hydrogen bonding penalty upon ligand binding has been shown to be a powerful strategy for the discovery of EphB4 inhibitors. In particular, compounds **1** and **2** showed IC_{50} values of 300 and 380 nM, and were predicted to be involved in four hydrogen bonds with the hinge region connecting the N terminal and C terminal lobes of the kinase domain.

Here, we confirmed the main features of the binding mode obtained by automatic docking using X-ray crystallography. As the formation of hydrogen bonds with Glu664 has been shown to be crucial for the design of highly potent kinase inhibitors, the addition of hydrogen bond donors was investigated. Chemical modifications of compounds **1** and **2** provided compound **15** with low nanomolar inhibition activity in enzymatic and cellular phosphorylation assay.

Compound **15** exhibited significant fluorescence properties (λ_{abs} = 370 nm, λ_{em} = 430 nm, φ = 0.25), making it suitable for *in vitro* imaging. Fluorescence emission was detected within the cellular cytoplasm of MDA-MB-231 cells after 2 hours of incubation, and significant changes in morphology were observed, such as cell shrinkage and disruption of cellular adhesion. Compound **15** appears as a promising EphB4 inhibitor with high potency and remarkable cell membrane permeation. The cellular activity of **15** on additional cell lines is now being pursued extensively.

4.10 Experimental section

4.10.1 Chemistry

All reactions, unless otherwise stated, were carried out under a nitrogen atmosphere using standard Schlenk-techniques. All reagents were used as received unless otherwise noted. Solvents were purchased in the best quality available, degassed by purging thoroughly with nitrogen and dried over activated molecular sieves of appropriate size. Alternatively, they were purged with argon and passed through alumina columns in a solvent purification system (Innovative Technology). Reactions were monitored by thin layer chromatography (TLC) using Merck TLC silica gel 60 F_{254}. Flash column chromatography was performed over silica gel (230-400 mesh). NMR spectra were recorded on AV2 400 or AV2 500 MHz Bruker spectrometers. Chemical shifts are given in ppm. The spectra are calibrated to the residual ^{1}H and ^{13}C signals of the solvents. Multiplicities are abbreviated as follows: singlet (s), doublet (d), triplet (t), quartet (q), doublet-doublet (dd), quintet (quint), septet (sept), multiplet (m), and broad (br). Melting points were determined on a Büchi Melting Point B-540 instrument.

High-resolution electrospray ionization mass spectrometry was performed on a Finnigan MAT 900 (Thermo Finnigan, San Jose, CA, USA) double-focusing magnetic sector mass spectrometer. Ten spectra were acquired. A mass accuracy ≤2 ppm was obtained in the peak matching acquisition mode by using a solution containing 2 µL PEG200, 2 µL PPG450, and 1.5 mg NaOAc (all obtained from Sigma-Aldrich, Buchs, Switzerland) dissolved in 100 mL MeOH (HPLC Supra grade, Scharlau, E-Barcelona) as internal standard.

2-(3-Chloroquinoxalin-2(1H)-ylidene)malononitrile (4a)

To a solution of NaH (250 mg, 10.45 mmol) in anhydrous dimethoxyethane (20 mL) was carrefully added malonitrile (663 mg, 10.04 mmol). The reaction mixture was stirred at 25 °C for 30 min and 2,3-dichloroquinoxaline (**3**, 1 g, 5.02 mmol) was added. The resulting mixture was stirred at 25 °C for 3 h and then heated to reflux for one additional hour. The reaction mixture was cooled to 0 °C and 0.4 M HCl (30 mL) was slowly added. After stirring for 30 min at 0 °C, the formed precipitate was filtered off and recrystallized in EtOH to afford the desired product in pure form (845 mg, 74% yield). ^{1}H NMR (400 MHz, DMSO-d_6): δ = 7.67-7.62 (m, 2H), 7.57 (ddd, *J* = 8.4 Hz, *J* = 6.8 Hz, *J* = 1.1 Hz, 1H), 7.36 (ddd, *J* = 8.1 Hz, *J* = 6.8 Hz, *J* = 1.2 Hz, 1H), NH not observed; ^{13}C NMR (100 MHz, DMSO-d_6): δ = 149.3, 141.9, 134.8, 134.1, 131.3, 127.5, 125.5, 120.2, 118.6, 44.5, one C missing due to overlapping; IR (film): $\tilde{\upsilon}$ = 3235, 3203, 3100, 3056, 3013, 2979, 2220, 2208, 1616, 1577, 1488, 1410, 1096, 969, 799, 764, 594 cm^{-1}; MS (ESI): *m/z*: calcd for $C_{11}H_5ClN_4Na^+$: 251.0, found: 250.9.

General procedure A for the cyclization of 2-(3-chloroquinoxalin-2-yl)malononitrile

To a mixture of 2-(3-chloroquinoxalin-2-yl)malononitrile (**4a**, 1 eq) in a 1 : 1 EtOH – Toluene solution (0.09 M) was added the primary amine (4 eq). The reaction mixture was heated to 130 °C for 3 h, cooled, and the formed solid was filtered off and washed with EtOH to afford the corresponding products in pure form. This method was used to obtain compounds **5**-**7**.

General procedure B for the cyclization of 2-(3-chloroquinoxalin-2-yl)malononitrile

To a mixture of 2-(3-chloroquinoxalin-2-yl)malononitrile (**4a**, 1 eq) in a 1 : 1 EtOH – Toluene solution (0.09 M) was added the primary amine (4 eq). The reaction mixture was heated to 160 °C for 3 h, concentrated under reduced pressure and purified by flash chromatography on silica gel. The carbonitrile was dissolved in sulfuric acid (0.15 M) and the solution was stirred for 30 min at 25 °C. The reaction was poured in ice cold water, basified with a solution of ammonium hydroxide (25%) and extracted with EtOAc. The organic layer was dried over $MgSO_4$, filtered, and concentrated under reduced pressure. Purification by column chromatography on silica gel afforded the desired compounds in pure form. This method was used to obtain intermediates **8** and **9** and final products **13** and **14**.

General procedure for the hydrolysis of carbonitriles

A solution of the carbonitrile (1 eq) in sulfuric acid (0.15 M) was stirred for 30 min at 25 °C. The reaction was poured in ice cold water, basified with a solution of ammonium hydroxide (25%) and the formed solid was filtered off. Extraction with EtOAc and a saturated solution of $NaHCO_3$ afforded the corresponding products in pure form. This method was used to obtain compounds **10**, **11** and **15**.

Preparation of compound 12

A solution of **7** (186 mg, 0.56 mmol) in sulfuric acid (3 mL) was stirred for 30 min at 25 °C. The reaction was poured in ice cold water and the formed solid was filtered off and washed with water and hexane. The crude product was then heated in acetone (5 mL) at 160 °C for 1 h. After cooling to room temperature, the solid was filtered off and washed with hexane to afford the desired product in pure form (143 mg, 72% yield).

4.10.2 FRET based enzymatic assay

Compounds were tested in the Z'-LYTE^{TM} Kinase Assay Kit–Tyr 1 Peptide (Invitrogen, USA) in a Corning 384 well microtiter plate. Fluorescence progress curves were measured upon excitation at 400 nm and emission at 445 and 520 nm. The assay contained a final concentration of EphB4 and ATP of 25 ng/µL and 125 µM (which is near its K_m), respectively, and was run at room temperature for 2 h. IC_{50} values (inhibitor concentration at which enzyme activity is reduced by 50%) are determined after carrying out assays at ten different concentrations between 20 µM and 10 pM.

4.10.3 Phage display based binding assay

The experiments were performed at Ambit Biosciences Corporation using binding assays as previously described.[24] Briefly, kinases were expressed as fusion proteins to T7 phage. In general, full-length constructs were used for small kinases and catalytic domains for large kinases. T7-kinase-tagged phage strains were mixed with known kinase inhibitors immobilized on streptavidin-coated magnetic beads and with test compounds at a single concentration of 1 µM. Test compounds that bind to the kinase ATP site displace the immobilized ligand from the kinase/phage, which is detected using quantitative PCR. The results are reported as the percentage of kinase/phage remaining bound to the ligand/beads, relative to a control (DMSO lacking a test compound). High affinity compounds have %control = 0, while weaker binders have higher %control values.

4.10.4 Fluorescence measurements

Absorption and fluorescence measurements were recorded on a SpectraMax M5 using a quartz cuvette (1 cm). A 5 mM solution of inhibitor (in 100% DMSO) was diluted in acetonitrile and its optical density at 370 nm was adjusted to 0.075 ± 0.01. Fluorescence was measured upon excitation at 370 nm. Quantum yields were calculated using quinine hemisulfate monohydrate (φ_R = 0.53) in 1 M H_2SO_4 as a reference and according to:

$$\varphi = \varphi_R \frac{F}{F_R} \frac{A_R}{A} \frac{n^2}{n_R^2}$$

φ_R corresponds to the quantum yield of the reference. n and n_R correspond to the refractive index of the solvents used for the inhibitor and for the reference. A and A_R are the optical densities of the inhibitor and of the reference. F and F_R correspond to the integrated emission spectra.

4.10.5 Biological assays

Cellular phosphorylation assays

The following experiments were performed at ProQinase GmbH. Mouse embryonal fibrablast cells, stably transfected to overexpress full-length human EphB4, were plated at 4×10^4 cells/well in DMEM supplemented with 10% FCS in 48-well culture dishes. Medium was replaced by DMEM without FCS before test compounds prediluted in 100% DMSO were added (final DMSO concentration of 1%). After incubation for 90 min at 37 °C, cells were stimulated for 2 h at 4 °C using murine ephrinB2-Fc at a final concentration of 2 µg/ml. Quantification of EphB4 phosphorylation was assessed in a 96-well plate via sandwich ELISA using a myc capture antibody and an antiphosphotyrosine detection antibody.

Cellular proliferation assay

MDA-MB-231 cells were maintained at 37 °C, 5% CO_2 atmosphere in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% (v/v) fetal bovine serum, 2 mM L-glutamine, 100 units/mL penicillin and 100 mg/mL streptomycin. Cells were plated at 10,000 cells per well in 96-well culture dishes and allowed to incubate for 24 h. A 5 mM solution of inhibitor (in 100% DMSO) was serially diluted in the culture media, added to the cells, and allowed to incubate for 24 h. Control cells were treated with the same DMSO concentrations. A solution of resazurin (0.86 M in PBS) was diluted 10 times in the culture media, added in each well and incubated for 3 additional hours. Fluorescence was measured upon excitation at 560 nm and emission at 590 nm. Resazurin becomes fluorescent under mitochondrial reduction and its emission directly correlates with the number of living cells. IC_{50} values were determined by plotting the fraction of surviving cells against the log of the drug concentration.

Confocal microscopy

MDA-MB-231 cells were plated at 20,000 cells per well in height-well Lab-Tek chamber slides and allowed to incubate overnight. A 5 mM solution of **15** (in 100% DMSO) was serially diluted in the culture media and added to the cells. After 24 h of incubations, cells were washed with PBS and fixed with 4% formaldehyde for 10 minutes at room temperature. Actin filaments were stained for 1 h using a 0.1 µM solution of tetramethylrhodamine isothiocyanate (TRITC)-phalloidin. After washing with PBS, images were taken using a 40 × 1.25 oil objective in a Leica wide field microscope.

4.10.6 X-ray crystallography

Protein expression and purification

A clone of the EphA3 kinase domain (residues: 606-947) was obtained from Prof. Sirano Dhe-Paganon's group[25] and expressed in *Escherichia coli* strain BL21 (DE3). Cells expressing EphA3 were induced with a 1 mM solution of isopropyl-beta-D-thiogalactopyranoside (IPTG) for 12 h at 15 °C. Cell pellets were resuspended in buffer A (50 mM Tris, pH 8.0, and 100 mM NaCl, supplemented with protease inhibitors) and lysed by sonication. After centrifugation at 15,000 rpm for 1 h, the soluble fraction of EphA3 was purified using HisTrap FF crude and HiTrap Q HP columns (GE Healthcare), followed by gel filtration chromatography (Superdex75; GE Healthcare). The appropriate fractions were combined and concentrated to ~10 mg/mL using Amicon filter devices (10 kDa as cutoff) in a storage solution (100 mM sodium chloride and 10 mM Tris-HCl pH 8.0, 5% glycerol). The resulting solution was aliquoted and stored at -80 °C for further usage.

Crystallization, Data Collection, and Structure Determination

Crystals of the EphA3 kinase domain were grown at 20 °C using the hanging drop vapor diffusion method. Equal volumes of protein and reservoir solutions (0.1 M sodium cacodylate pH 6.5, 0.15 M

ammonium sulfate, 22.5% PEG 3350) were mixed and crystals appeared after 1 to 2 days. A 5 mM solution of inhibitor (in 100% DMSO) was added into the hanging drop to reach a final DMSO concentration of 10% (v/v). The crystals were soaked for 1 to 24 h and flash-frozen in liquid nitrogen without extra cryoprotectant.

Refinement and validation

Data sets were collected on a MarCCD detector at the Swiss Light Source beamline XS06DA of the Paul Scherrer Institute (Villigen, Switzerland) and indexed, integrated and scaled with the XDS[26] and CCP4 programs.[27] The structures were solved by molecular replacement with PHASER[28] using the apo EphA3 kinase domain structure (PDB entry 2GSF) as a search model and refined with PHENIX.[29]

4.11 Acknowledgement

We thank the staff of beamline PX of the Swiss Light Source (Villingen, Switzerland) for a valuable technical assistance. In particular, we would like to thank Dr. Xiao-Dan Li and Chitra Rajendran for their support.

4.12 References

(1) Zhao, H.; Huang, D. Hydrogen bonding penalty upon ligand binding. *PLoS One* **2011**, *6*, e19923.

(2) Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. C. Green fluorescent protein as a marker for gene expression. *Science* **1994**, *263*, 802-805.

(3) Tsien, R. Y. The green fluorescent protein. *Annu. Rev. Biochem.* **1998**, *67*, 509-544.

(4) Son, J.; Lee, J. J.; Lee, J. S.; Schuller, A.; Chang, Y. T. Isozyme-specific fluorescent inhibitor of glutathione s-transferase omega 1. *ACS Chem. Biol.* **2010**, *5*, 449-453.

(5) Miller, L. W.; Sable, J.; Goelet, P.; Sheetz, M. P.; Cornish, V. W. Methotrexate conjugates: a molecular *in vivo* protein tag. *Angew. Chem. Int. Ed.* **2004**, *43*, 1672-1675.

(6) Kim, D.; Lee, H.; Jun, H.; Hong, S. S.; Hong, S. Fluorescent phosphoinositide 3-kinase inhibitors suitable for monitoring of intracellular distribution. *Bioorg. Med. Chem.* **2011**, *19*, 2508-2516.

(7) Yenugonda, V. M.; Deb, T. B.; Grindrod, S. C.; Dakshanamurthy, S.; Yang, Y.; Paige, M.; Brown, M. L. Fluorescent cyclin-dependent kinase inhibitors block the proliferation of human breast cancer cells. *Bioorg. Med. Chem.* **2011**, *19*, 2714-2725.

(8) Uddin, M. J.; Crews, B. C.; Blobaum, A. L.; Kingsley, P. J.; Gorden, D. L.; McIntyre, J. O.; Matrisian, L. M.; Subbaramaiah, K.; Dannenberg, A. J.; Piston, D. W.; Marnett, L. J. Selective visualization of cyclooxygenase-2 in inflammation and cancer by targeted fluorescent imaging agents. *Cancer Res.* **2010**, *70*, 3618-3627.

(9) Bae, Y.; Nishiyama, N.; Fukushima, S.; Koyama, H.; Yasuhiro, M.; Kataoka, K. Preparation and biological characterization of polymeric micelle drug carriers with intracellular pH-triggered drug release property: tumor permeability, controlled subcellular drug distribution, and enhanced *in vivo* antitumor efficacy. *Bioconjugate Chem.* **2005**, *16*, 122-130.

(10) Klein, A. V.; Hambley, T. W. Platinum drug distribution in cancer cells and tumors. *Chem. Rev.* **2009**, *109*, 4911-4920.

(11) Bamborough, P.; Angell, R. M.; Bhamra, I.; Brown, D.; Bull, J.; Christopher, J. A.; Cooper, A. W.; Fazal, L. H.; Giordano, I.; Hind, L.; Patel, V. K.; Ranshaw, L. E.; Sims, M. J.; Skone, P. A.; Smith, K. J.; Vickerstaff, E.; Washington, M. N-4-Pyrimidinyl-1H-indazol-4-amine inhibitors of Lck: indazoles as phenol isosteres with improved pharmacokinetics. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 4363-4368.

(12) Lafleur, K.; Huang, D.; Zhou, T.; Caflisch, A.; Nevado, C. Structure-based optimization of potent and selective inhibitors of the tyrosine kinase erythropoietin producing human hepatocellular carcinoma receptor B4 (EphB4). *J. Med. Chem.* **2009**, *52*, 6433-6446.

(13) Bardelle, C.; Barlaam, B.; Brooks, N.; Coleman, T.; Cross, D.; Ducray, R.; Green, I.; Brempt, C. L.; Olivier, A.; Read, J. Inhibitors of the tyrosine kinase EphB4. Part 3: identification of non-benzodioxole-based kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 6242-6245.

(14) Barlaam, B.; Ducray, R.; Lambert-van der Brempt, C.; Ple, P.; Bardelle, C.; Brooks, N.; Coleman, T.; Cross, D.; Kettle, J. G.; Read, J. Inhibitors of the tyrosine kinase EphB4. Part 4: Discovery and optimization of a benzylic alcohol series. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 2207-2211.

(15) Zuccotto, F.; Ardini, E.; Casale, E.; Angiolini, M. Through the "gatekeeper door": exploiting the active kinase conformation. *J. Med. Chem.* **2010**, *53*, 2681-2694.

(16) Ishikawa, M.; Hashimoto, Y. Improvement in aqueous solubility in small molecule drug discovery programs by disruption of molecular planarity and symmetry. *J. Med. Chem.* **2011**, *54*, 1539-1554.

(17) Pratt, E. F.; Kereszte.Jc Syntheses of Indolizino- and Dihydroindolizinoquinoxalines. *J. Org. Chem.* **1967**, *32*, 49-53.

(18) Obafemi, C. A.; Pfleiderer, W. Synthesis and some reactions of 3-chloro-2-(cyanomethylene)-1,2-dihydroquinoxalines. *Molecules* **2004**, *9*, 223-231.

(19) Otomasu, H.; Ohmiya, S.; Sekuguch.T; Takahash.H Synthesis of Condensed Quinoxalines .2. A New Synthesis of Pyrrolo- 2,3-B!Quinoxalines. *Chem. Pharm. Bull.* **1970**, *18*, 2065-2069.

(20) Lakowicz, J. R., *Principles of fluorescence spectroscopy*. Second ed. ed.; 1999.

(21) Lombardo, L. J.; Lee, F. Y.; Chen, P.; Norris, D.; Barrish, J. C.; Behnia, K.; Castaneda, S.; Cornelius, L. A.; Das, J.; Doweyko, A. M.; Fairchild, C.; Hunt, J. T.; Inigo, I.; Johnston, K.; Kamath, A.; Kan, D.; Klei, H.;

Marathe, P.; Pang, S.; Peterson, R.; Pitt, S.; Schieven, G. L.; Schmidt, R. J.; Tokarski, J.; Wen, M. L.; Wityak, J.; Borzilleri, R. M. Discovery of N-(2-chloro-6-methyl- phenyl)-2-(6-(4-(2-hydroxyethyl)- piperazin-1-yl)-2-methylpyrimidin-4- ylamino)thiazole-5-carboxamide (BMS-354825), a dual Src/Abl kinase inhibitor with potent antitumor activity in preclinical assays. *J. Med. Chem.* **2004**, *47*, 6658-6661.

(22) Fox, B. P.; Kandpal, R. P. Invasiveness of breast carcinoma cells and transcript profile: Eph receptors and ephrin ligands as molecular markers of potential diagnostic and prognostic application. *Biochem. Biophys. Res. Commun.* **2004**, *318*, 882-892.

(23) Mgbonyebi, O. P.; Russo, J.; Russo, I. H. Roscovitine induces cell death and morphological changes indicative of apoptosis in MDA-MB-231 breast cancer cells. *Cancer Res.* **1999**, *59*, 1903-1910.

(24) Karaman, M. W.; Herrgard, S.; Treiber, D. K.; Gallant, P.; Atteridge, C. E.; Campbell, B. T.; Chan, K. W.; Ciceri, P.; Davis, M. I.; Edeen, P. T.; Faraoni, R.; Floyd, M.; Hunt, J. P.; Lockhart, D. J.; Milanov, Z. V.; Morrison, M. J.; Pallares, G.; Patel, H. K.; Pritchard, S.; Wodicka, L. M.; Zarrinkar, P. P. A quantitative analysis of kinase inhibitor selectivity. *Nat Biotechnol* **2008**, *26*, 127-132.

(25) Choi, Y.; Syeda, F.; Walker, J. R.; Finerty, P. J., Jr.; Cuerrier, D.; Wojciechowski, A.; Liu, Q.; Dhe-Paganon, S.; Gray, N. S. Discovery and structural analysis of Eph receptor tyrosine kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 4467-4470.

(26) Kabsch, W. Automatic Processing of Rotation Diffraction Data from Crystals of Initially Unknown Symmetry and Cell Constants. *J. Appl. Crystallogr.* **1993**, *26*, 795-800.

(27) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **1994**, *50*, 760-763.

(28) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40*, 658-674.

(29) Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L. W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **2002**, *58*, 1948-1954.

Chapter 5: Addendum - Complementing ultrafast shape recognition with an optical isomerism descriptor

5 Addendum - Complementing ultrafast shape recognition with an optical isomerism descriptor

5.1 Introduction

Shape complementarity is essential in macromolecular recognition and binding of small molecules to proteins because of the sensitivity of the van der Waals energy at separations close to the optimal distance. There is abundant experimental evidence that small molecules with shape similar to known active compounds are likely to have similar biological activities.[1] Therefore, screening of databases of three-dimensional (3D) molecular structures can be performed by comparison of molecular shapes.[2-4] Several methods have been developed and applied in the past few decades to identify compounds similar to a query molecule.[5-10] They are useful whenever one or more inhibitors of a target protein are known particularly when the 3D structure of the protein is not available.

Recently, a method termed Ultrafast Shape Recognition (USR) has been developed for searching very large databases of molecular structures.[11] Despite its recent publication, USR has already been used in several drug design projects[3,12-15] because of its simplicity and efficiency. Importantly, the molecules do not need to be superposed. Only, the distances between each atom of the molecule and four molecular locations are calculated for USR: the molecular centroid (ctd), the atom closest to ctd (cst), the atom farthest to ctd (fct), and the atom farthest to fct (ftf). The shape of a molecule is then encoded by 12 descriptors independent of the number of atoms. The first descriptor is the mean of atomic distances from ctd $\mu_1^{ctd} \equiv \frac{1}{N}\sum_{j=1}^{N} d_j^{ctd}$, where d_j is the distance of the j^{th} atom from ctd, and N is the number of atoms in the molecule. The second descriptor is the square root of the second central moment of the distribution of the same atomic distances.

$$\mu_2^{ctd} \equiv \frac{1}{N}\left[\sum_{j=1}^{N}(d_j^{ctd} - \mu_1^{ctd})^2\right]^{1/2}$$

The third descriptor is the cubic root of the third central moment of the same distribution:

$$\mu_3^{ctd} \equiv \frac{1}{N}\left[\sum_{j=1}^{N}(d_j^{ctd} - \mu_1^{ctd})^3\right]^{1/3}$$

which is a measure of asymmetry. The remaining nine descriptors are calculated analogously using cst, fct, and ftf. Since only intramolecular distances are used in the 12 descriptors, the USR is not able to distinguish mirror images.

Here, we supplement the original USR method[11] with an optical isomerism descriptor that is able to discriminate a molecule from its mirror image, and is therefore particularly useful for clustering conformers and searching 3D databases. Our extension of USR (called USR:OptIso) is first tested on

three pairs of conformations of kinase inhibitors and 15 pairs of different types of isomers. Then similarity scores based on USR and USR:OptIso (non-superposition methods) for 1.6×10^{10} pairs of conformers of 2.7 millions small molecules are compared with the ones based on Gaussian molecular volume overlap[16] calculated by ROCS (superposition method, OpenEye Scientific Software).

5.2 Methods

5.2.1 Optical isomerism descriptor

Considerable efforts have been devoted to symmetry detection in chemistry and chemoinformatics. In particular, several methods have been developed to analyze chirality. These include two-dimensional descriptors[17-19] for the prediction of the major product of stereoselective reactions,[20-22] and three-dimensional descriptors (chiral topological indices) as complement to distance matrices in quantitative stereochemical structure–activity relationship models.[23-29]

Here, the following vectors are introduced for the efficient evaluation of the optical isomerism descriptor: **a**=**cst**−**ctd**, **b**=**fct**−**ctd**, and **c**=**ftf**−**ctd**, where **ctd**, **cst**, **fct**, and **ftf** are the vectors connecting the origin of the coordinates to each of the four molecular locations (Figure 1). The optical isomerism descriptor is defined as the cubic root of the scalar triple product (or mixed product) of **a**, **b**, and **c**, i.e., optical isomerism descriptor=$[\mathbf{c}.(\mathbf{a}\times\mathbf{b})]^{1/3}$. The cubic root is used to obtain a unit of length (Å) as for the other 12 descriptors. The computational cost for evaluating the optical isomerism descriptor is neglectable, since the coordinates of the four molecular locations have already been calculated for the other 12 descriptors.

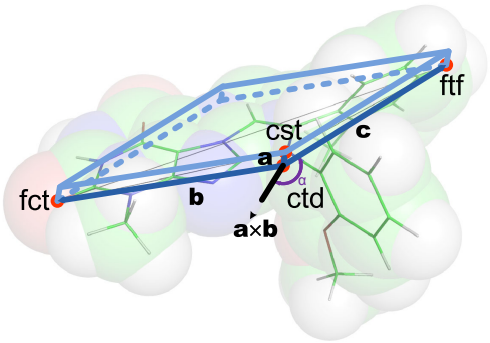


Figure 1. Optical isomerism descriptor. The four molecular locations of a conformer of compound **1** are denoted with red circles, and the three vectors **a**, **b**, and **c** with blue arrows. The optical isomerism descriptor is the cubic root of the volume of the parallelepiped with blue edges. The sign of the optical isomerism descriptor is negative for this conformer because **c** and **a**×**b** form an obtuse angle (violet angle α). The mirror image of this conformer has a positive value of the optical isomerism descriptor, and is shown in Figure 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

During the writing of this manuscript, Armstrong et al. reported a modification of USR that is able to distinguish enantiomers.[30] They use the cross product of two vectors spanning three of the four USR molecular locations (ctd, fct, and ftf) to define a fourth location which is different from cst. In contrast, the crucial component of our descriptor is the triple product of three vectors spanning all of the four locations. Moreover, Armstrong and collaborators replace three of the 12 USR descriptors (those involving cst) whereas we supplement the USR with the optical isomerism descriptor. Note that cst might contain useful information as Ballester et al. reported that the first moment of the distribution of distances from cst had a similar value across active molecules (Figure 3 in Ref. 3). In general, the atomic distances to ctd are different from that to cst as the separation between ctd and cst is usually between 0.4 and 2.0 Å for small molecules (see section 5.6). Since our optical isomerism descriptor contains more geometrical information than the one of Armstrong et al., it is expected to perform at least equivalently.

5.2.2 Similarity score

The inverse Manhattan distance is used as similarity score:[11]

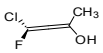
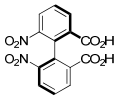
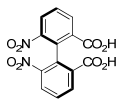
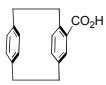
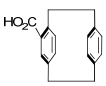
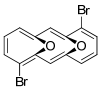
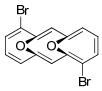
$$S_{pq} = \frac{1\ lenght\ unit}{1\ lenght\ unit + \frac{1}{n}\sum_{i=1}^{n}|M_i^p - M_i^q|}$$

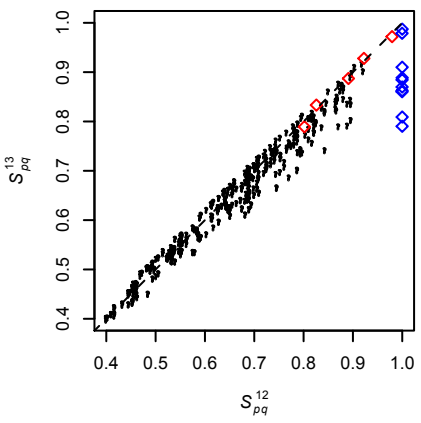
where M_i^p is the ith descriptor of the conformation p. Note that all the USR descriptors have a unit of length, which is Å here, therefore the S_{pq} is dimensionless. The addition of 1 length unit at the denominator prevents a division by zero in the case of identical 3D structures, and yields a similarity score of 1 for them. The similarity scores with the optical isomerism descriptor (S_{pq}^{13}) and without (S_{pq}^{12}) are compared in the next section.

5.3 Results and discussions

5.3.1 Discriminatory power of USR:OptIso and usefulness for clustering

The ability of the optical isomerism descriptor to discriminate isomers is presented in Figure 2. For a molecule and its mirror image the 12 descriptors of the original USR method are identical because they only depend on distance distributions. However, their optical isomerism descriptors are opposite because their four molecular locations are mirror images as well. In contrast, structural isomers, diastereoisomers, and other types of conformers that are not mirror image of each other have different distance distributions. Therefore, the first 12 descriptors are enough to discriminate them.

Isomer1	Isomer2	Type of Isomerism	S_{pq}^{12}	S_{pq}^{13}	OID1	OID2	ROCS
		Structural isomerism: functional group isomerism	0.891	0.887	0.153	-0.032	0.999
		Structural isomerism: position isomerism	0.802	0.789	0.396	-0.106	0.825
		Structural isomerism: skeletal isomerism	0.826	0.833	0.838	0.911	0.892
		Conformational isomerism	1.000	0.889	0.815	-0.815	0.667
		Diastereoisomerism: cis/trans	0.979	0.972	0.456	0.334	0.685
		Diastereoisomerism: E/Z	0.922	0.928	0.000	0.000	0.302
		Chirality with one stereogenic centers: R/S	1.000	0.910	0.642	-0.642	0.997
		Chirality without stereogenic centers: allenes	1.000	0.884	0.849	-0.849	0.997
		Chirality without stereogenic centers: alkylidenecycloalkanes	1.000	0.860	1.054	-1.054	0.953
		Chirality without stereogenic centers: spiranes	1.000	0.987	-0.087	0.087	0.963
		Chirality without stereogenic centers: byphenyls - atroposomerism	1.000	0.964	-0.238	0.238	0.993
		Chirality without stereogenic centers: helicenes	1.000	0.790	1.725	-1.725	0.703
		Chirality without stereogenic centers - planar chirality: cyclophanes	1.000	0.759	2.053	-2.053	0.322
		Chirality without stereogenic centers - planar chirality: annulenes	1.000	0.863	1.035	-1.035	0.947
		Chirality without stereogenic centers - planar chirality: cyclalkenes	1.000	0.809	-1.514	1.514	0.957



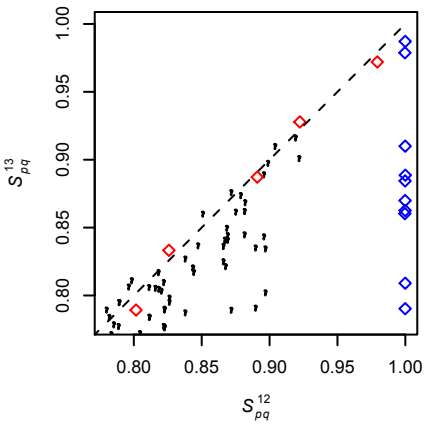


Figure 2. Application of the optical isomerism descriptors to isomers. All pairs of isomers can be distinguished by the USR:OptIso (13 descriptors) because the optical isomerism descriptors (OID1 and OID2) of mirror image isomers have opposite signs. The scatter plots show the pairwise USR comparisons of these 30 compounds (bottom left, full data set; bottom right, zoom-in on values close to 1). The original USR (S_{pq}^{12}) and USR:OptIso (S_{pq}^{13}) methods assign close similarities except for 10 pairs (blue diamonds) out of 435. The five pairs of isomers that can be distinguished by both S_{pq}^{12} and S_{pq}^{13} are denoted with red diamonds. The dashed line represents S_{pq}^{12} = S_{pq}^{13}. The last column of the table is the ROCS shape Tanimoto.[16]

A recently published inhibitor of the receptor tyrosine kinase Ephrin type-B receptor 4 (EphB4)[13] is used to illustrate the usefulness of the optical isomerism descriptor (Figure 3). The similarity score S_{pq}^{13} is able to distinguish the two mirror image conformers of compound **1** because of the opposite sign of their optical isomerism descriptors. In contrast, the two conformers of **1** have identical 12 descriptors based on the original USR method.[11]

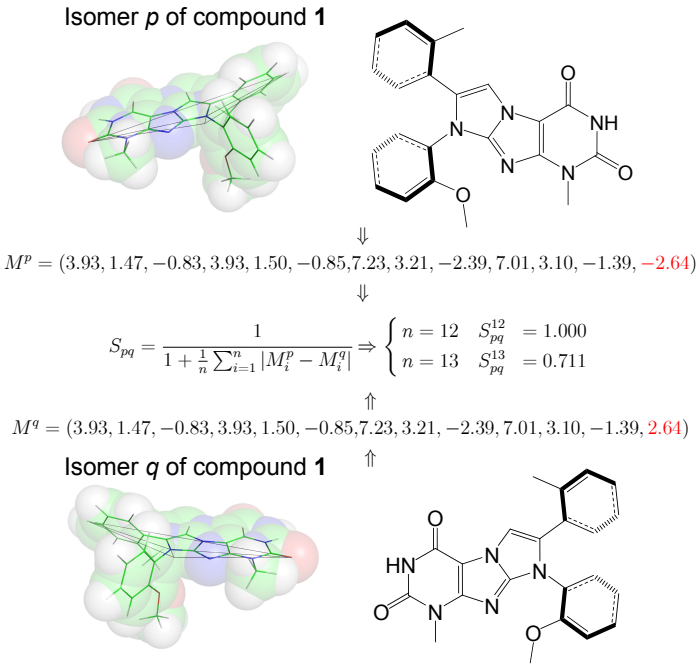


Figure 3. Two mirror images of the tyrosine kinase inhibitor **1**[13] generated by systematic bond rotation are distinguished by the optical isomerism descriptor. The four molecular locations are connected with black lines on the 3D structures which are shown with sticks and transparent CPK models. This example shows the high discriminating power of the optical isomerism descriptor whose usage results in a low similarity score S_{pq}^{13} = 0.711 for the two mirror images of **1** whereas they are not distinguished by the original USR method (S_{pq}^{12} = 1.000). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

The optical isomerism descriptor is useful for clustering as it can distinguish between different conformers/isomers that would be clustered together by the original USR. In our previous study,[13] multiple conformers of **1** were generated by systematic bond rotation and optimized to their nearest local minima using density functional theory. Two local minima were then considered identical if their similarity score S_{pq} was higher than 0.999. Interestingly, the opposite sign of the optical isomerism descriptor contributes significantly to the identification of mirror images (or pairs of conformers very close to mirror images), in particular when S_{pq}^{12} is close to 1 (Figure 4). Furthermore, the USR:OptIso was tested on two pairs of isomers of recently published kinase inhibitors (**2** and **3** in Figure 5).[31,32] The first 12 descriptors of USR have identical values, while the optical isomerism descriptor reduces the similarity score from 1 to 0.705 for compound **2** and from 1 to 0.650 for compound **3**.

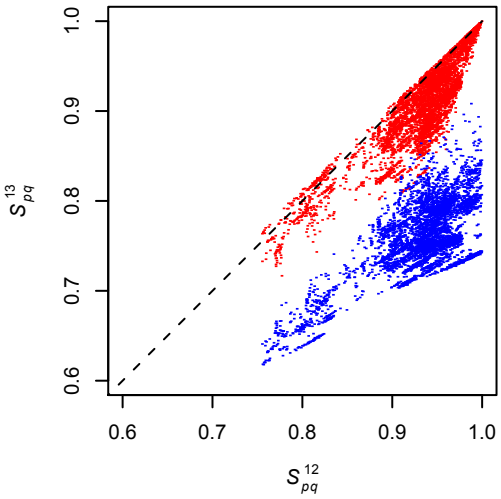


Figure 4. Scatter plot of the similarity scores with the optical isomerism descriptor (S_{pq}^{13}) and without (S_{pq}^{12}) for the 10,585 pairs of 146 local minima[13] of **1**. The color of each data point illustrates the sign of the optical isomerism descriptor (same and opposite signs are in gray and black, respectively). The dashed line represents $S_{pq}^{12} = S_{pq}^{13}$.

The optical isomerism descriptor can be used for searching (multi-)conformational libraries. As an example, using the similarity score S_{pq}^{13} yields only the conformer similar to the query whereas the conformations that are similar to its mirror image might be retrieved erroneously if one neglects the optical isomerism descriptor. This is a clear advantage of USR:OptIso with respect to the original USR.

Finally, it is necessary to verify that similar conformers of a given molecule yield very similar values of S_{pq}^{12} and S_{pq}^{13}. A set of 100 similar structures of the protein kinase inhibitor PP2[33] was used for assessing the robustness upon minor structural change of the original USR and USR:OptIso. A scatter plot is presented in section 5.6. This test indicates that slight changes in the coordinates yield minor changes in

both S_{pq}^{12} and S_{pq}^{13} when they are close to 1. Note that the high similarity range (i.e., values close to 1) is the most relevant case for virtual screening as only a small fraction of hits can be tested in practice.

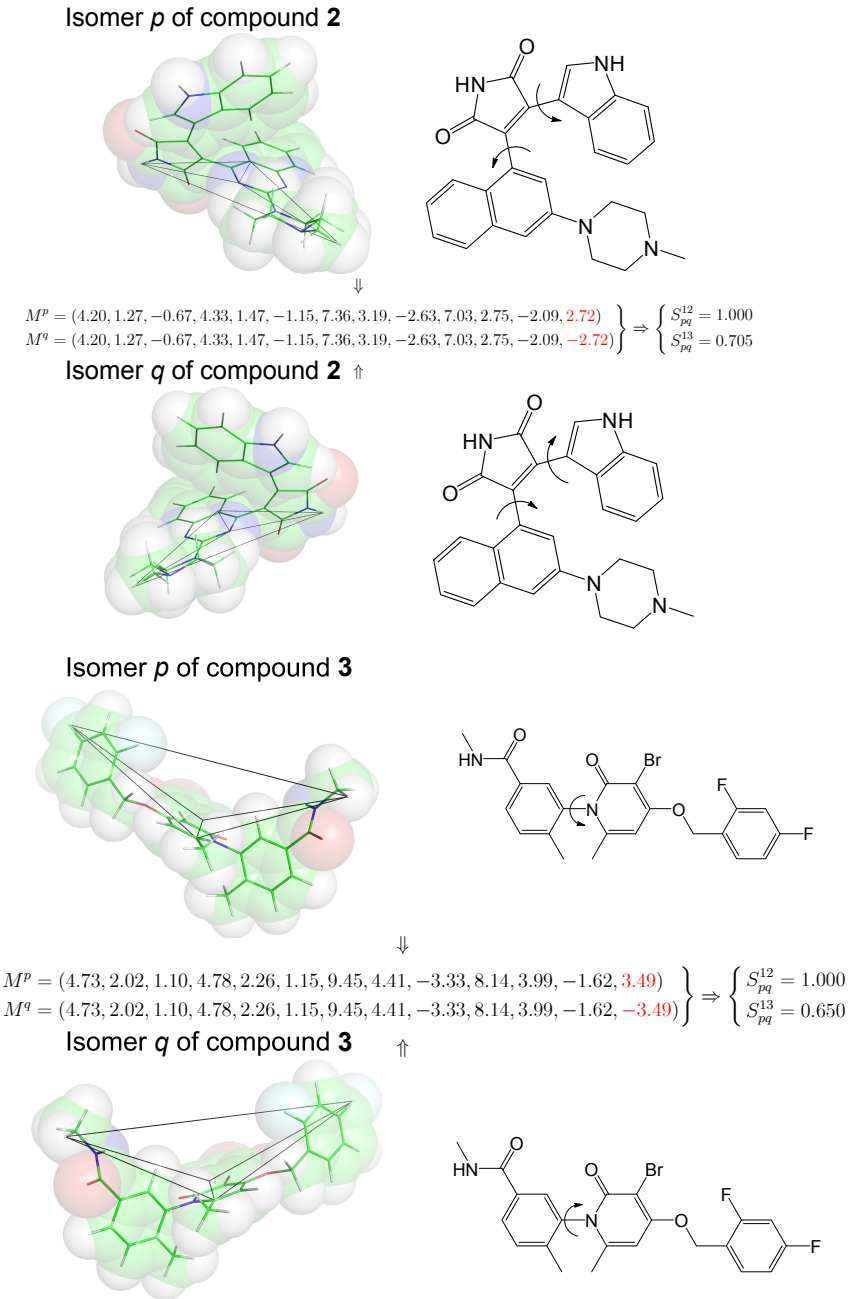


Figure 5. Discriminatory power of the optical isomerism descriptor. (a) The conformer of the protein kinase C inhibitor **2** observed in the X-ray structure[31] is shown in the top panel, while its mirror image is shown in the bottom panel. The similarity score decreases from 1 to 0.705 by taking into account the optical isomerism descriptor. (b) The active conformer of the p38 MAP kinase inhibitor **3**, shown in the top panel, is >100-fold more potent than its atropisomer shown in the bottom panel.[32] The similarity score decreases from 1 to 0.650 by taking into account the optical isomerism descriptor. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

5.3.2 Potential limitations of the optical isomerism descriptor

The optical isomerism descriptor is the cubic root of the (signed) volume of the parallelepiped defined by three vectors connecting four molecular locations. It is therefore equal to zero whenever the four molecular locations are coplanar. To estimate the frequency of the coplanarity of these four locations, we calculated the optical isomerism descriptor of 100,812,356 poses of about 2.7 million molecules (downloaded from the 2007 version of the ZINC library[34]) generated by high-throughput docking into EphB4.[35] Strikingly, only 28 poses (of 24 molecules) have an optical isomerism descriptor smaller than 0.01 in absolute value, which indicates that the optical isomerism descriptor is able in the vast majority of cases to clearly discriminate a given isomer from its mirror image. Some of these 28 poses are (close to) coplanar, while in others of them, the majority of the atomic nuclei are in a plane, and the nuclei out of the plane happen to be excluded (they are neither closest nor farthest to a centroid) from the four centroids. Note that a value of the optical isomerism descriptor different from zero does not necessarily imply that a molecule is chiral. For this reason, we prefer to use the term "optical isomerism" rather than "chirality" descriptor.

5.3.3 Comparison between similarity scores calculated by USR and ROCS

It is interesting to compare the USR-based similarity scores with a metric based on superimposed volume. The correlation coefficient between the similarity score based on Gaussian molecular shape overlap (ROCS shape Tanimoto[16]) and either S_{pq}^{12} or S_{pq}^{13} is 0.64 for the 435 possible pairs of the 30 compounds in Figure 2. This relatively low correlation is due to the fact that the similarity score evaluations are based on two different procedures. In the former, the similarity is calculated by volume overlap percentage after structural superposition, whereas in USR which does not require overlap, the similarity is evaluated using distributions of nuclei distances from the molecular centroids. Moreover, neither the iterative maximizing of the overlapped molecular volume in ROCS nor the maximum/minimum function for determining centroids in USR is continuous with respect to the Cartesian coordinates of the atomic nuclei, which are used as input.

For an in-depth comparison of similarity scores calculated by USR or USR:OptIso and ROCS shape Tanimoto the aforementioned 100 million poses of about 2.7 million molecules from ZINC were used to generate 1.6×10^{10} pairs of conformations. Notably, these two methods have very different algorithms to evaluate molecular shape similarity. Cartesian coordinates are the only input required to calculate USR scores, whereas ROCS also needs atomic radii to evaluate the Gaussian molecular volume. To compare the conformation pairs filtered by different similarity cutoffs, the "accuracy" was defined as the number of pairs for which both scores exceeded the cutoff divided by the pairs for which only the USR similarity

score exceeded the cutoff (i.e., C/(A + C) in Figure 6). In the same way, the "completeness" was defined as the ratio of the number of pairs for which both scores exceeded the cutoff to the pairs for which only ROCS shape Tanimotos exceeded the cutoff (i.e., C/(B + C) in Figure 6).

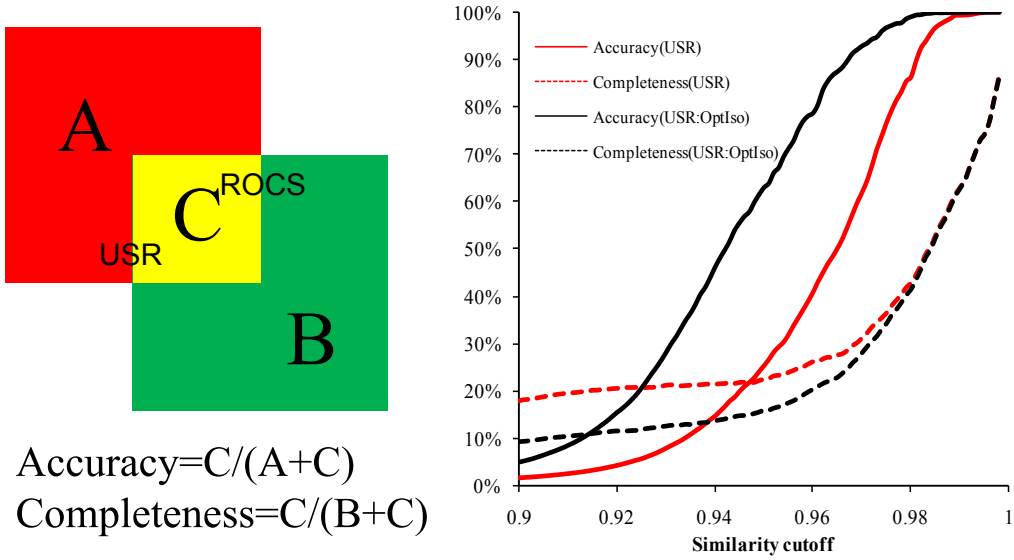


Figure 6. Accuracy and completeness of USR similarity score in reproducing ROCS shape Tanimoto. In the left panel, the red and the green squares are the sets of conformation pairs for which USR similarity scores and ROCS shape Tanimoto satisfy a particular cutoff, respectively. Their overlap (yellow part) represents the conformation pairs for which both similarity scores satisfy the cutoff. In the right panel, the cutoff varies from 0.9 to 1.0, which covers the range interesting for virtual screening applications. The solid and dashed lines represent accuracy and completeness, respectively, while the red and black color denote the USR and USR:OptIso results, respectively.

Both accuracy and completeness increase monotonously with the similarity cutoff (Figure 6). The accuracy of USR:OptIso has improved compared with the original USR, because the mirror images of the query conformation have been eliminated by the opposite optical isomerism descriptors. For instance, 91.25% of the conformation pairs that have USR:OptIso scores ≥0.968 also have ROCS shape Tanimotos ≥0.968, whereas the percentage decreases to 57.17% for the original USR. If the similarity score cutoff is set to 0.98, the accuracy of the USR:OptIso increases to 99.02% compared to 86.30% for the original one. Panel (a) of Figure 7 shows "the most inaccurate" example whose USR score is higher than the ROCS shape Tanimoto.

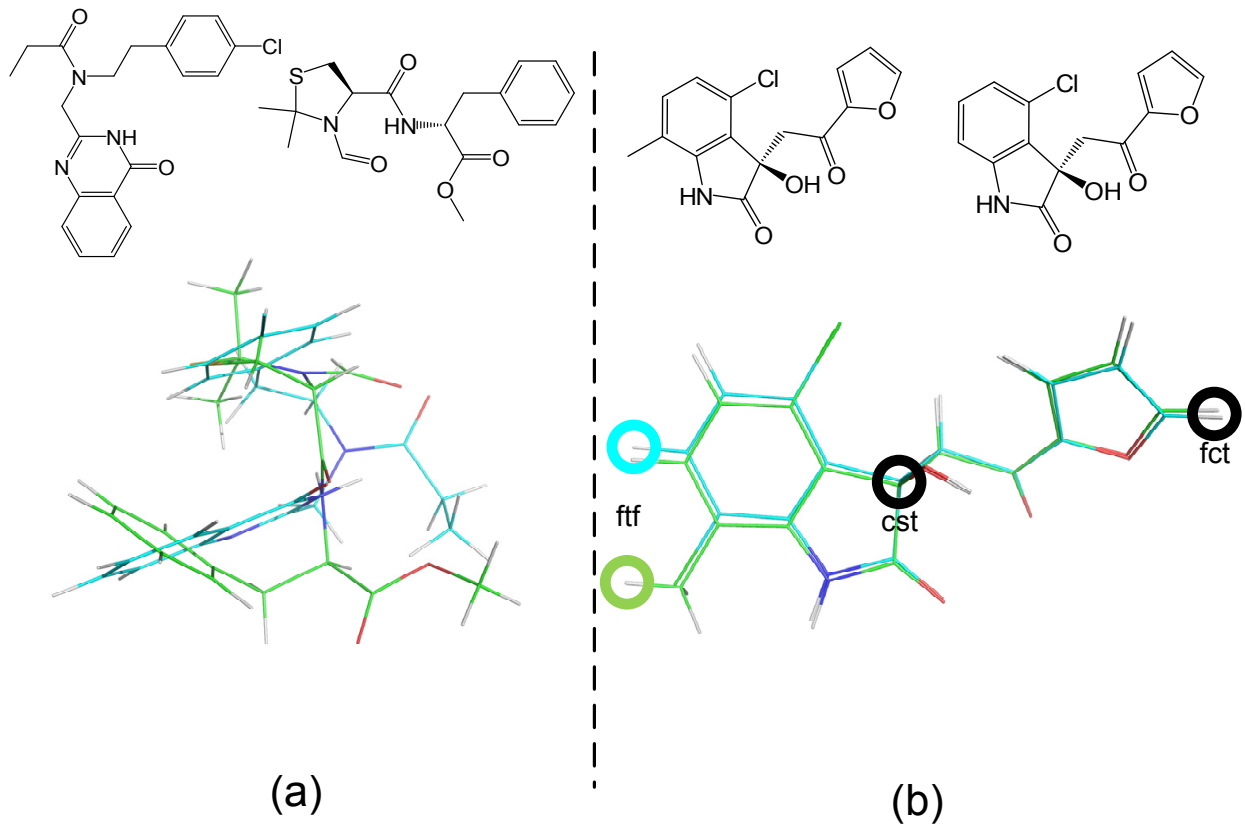


Figure 7. Examples of large discrepancies between USR and ROCS shape Tanimoto. (a) High USR similarity score but low ROCS shape Tanimoto. The overlaid conformations were optimized by ROCS. The USR and USR:OptIso similarity scores are 0.9807 and 0.9810. The ROCS shape Tanimoto is 0.676. This is the only case out of 1.6×10^{10} conformation pairs whose USR is larger than 0.98 and ROCS shape Tanimoto smaller than 0.7. (b) Low USR similarity score but high ROCS shape Tanimoto. The USR and USR:OptIso similarity scores are 0.6988 and 0.6981. The ROCS shape Tanimoto is 0.999. These two molecules are different in only one substituent. The additional methyl group in the green conformation changes the location of ftf from the phenyl hydrogen (of the cyan conformation) to the methyl hydrogen. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

The completeness of USR methods is low. One of the main reasons is that USR overestimates the difference of conformation pairs that are different at the extremity of the molecule where fct and/or ftf are defined. Figure 7(b) shows an example of a conformation pair whose ROCS shape Tanimoto is higher than the USR similarity score. In this example, USR overestimates the conformational difference because of a change of the ftf location. The similarity calculated by USR:OptIso is often lower than the one calculated by USR (Figure 4) due to the influence of the optical isomerism descriptor. Therefore, USR:OptIso has lower completeness than USR. This difference in completeness becomes smaller when the similarity cutoff approaches 1. The low completeness is not surprising as ROCS shape Tanimoto distinguishes atomic elements by different van der Waals radii whereas both USR and USR:OptIso treat

all atoms equally. Moreover, the optimal volume overlap has to be calculated in ROCS shape Tanimoto for every pair of conformations while the USR methods are significantly more efficient (more than 5000 times faster[11]) as they use only interatomic distances.

5.4 Conclusions

The optical isomerism descriptor (defined as the mixed product of the three vectors from the molecular centroid to the three molecular locations cst, fct, ftf) is an efficient and robust tool for shape comparison. It can be used as a supplement of the original 12 USR descriptors, which are based solely on distance distributions, while the optical isomerism descriptor is able to distinguish mirror images. It is therefore helpful for analyzing molecules with stereogenic centers, atropisomerism, and in the clustering of conformers generated by systematic bond rotation. Moreover, it can be used for the efficient search of molecular conformations that are superposable on the query structure. Finally, a comparison of the USR similarity score with the ROCS shape Tanimoto shows that both accuracy and completeness increase monotonously with the similarity score cutoff. The accuracy of the USR:OptIso similarity score is always higher than the one based on the original USR, and the completeness of USR:OptIso is close to the one of USR in high similarity ranges, which are relevant for virtual screening.

5.5 Acknowledgment

We thank Dr. Ballester for providing a detailed description of USR. We thank OpenEye Scientific Software for providing an academic license of ROCS. Source code: the source code for calculating the optical isomerism descriptor is available at http://code.google.com/p/usrchirality/.

5.6 Additional experimental data

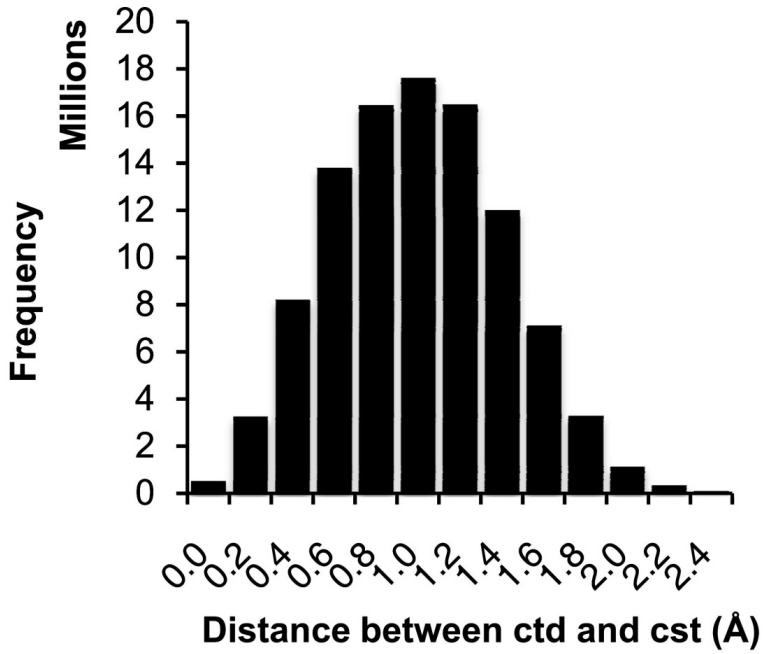


Figure 8. The distribution of ctd–cst distances of 100 million poses of 2.7 million of molecules. These molecules are from a subset of the ZINC library with molecular weight less than 500 Da.

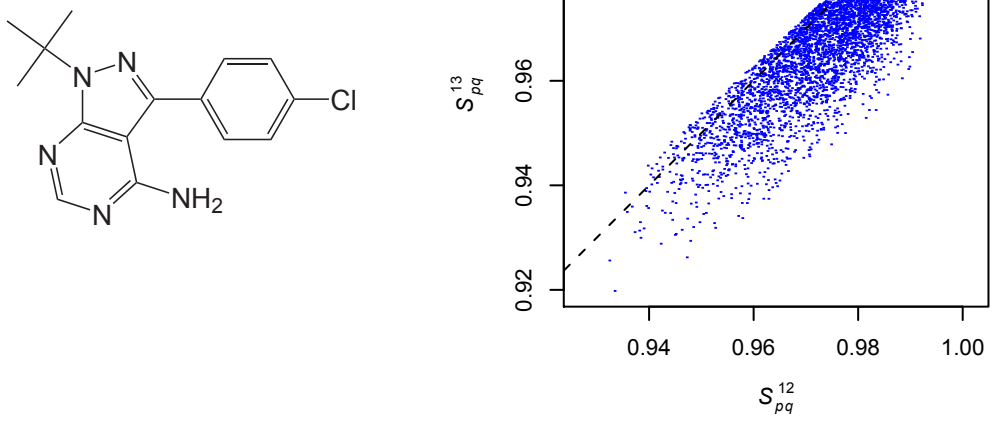


Figure 9. Pairwise comparison of 100 little fluctuated conformers of compound **4**. The dashed line represents $S_{pq}^{13} = S_{pq}^{12}$.

5.7 References

(1) McGaughey, G. B.; Sheridan, R. P.; Bayly, C. I.; Culberson, J. C.; Kreatsoulas, C.; Lindsley, S.; Maiorov, V.; Truchon, J. F.; Cornell, W. D. Comparison of topological, shape, and docking methods in virtual screening. *J. Chem. Inf. Model.* **2007**, *47*, 1504-1519.

(2) Rush, T. S.; Grant, J. A.; Mosyak, L.; Nicholls, A. A shape-based 3-D scaffold hopping method and its application to a bacterial protein-protein interaction. *J. Med. Chem.* **2005**, *48*, 1489-1495.

(3) Ballester, P. J.; Finn, P. W.; Richards, W. G. Ultrafast shape recognition: evaluating a new ligand-based virtual screening technology. *J. Mol. Graph. Model.* **2009**, *27*, 836-845.

(4) Kinnings, S. L.; Jackson, R. M. LigMatch: a multiple structure-based ligand matching method for 3D virtual screening. *J. Chem. Inf. Model.* **2009**, *49*, 2056-2066.

(5) Artymiuk, P. J.; Bath, P. A.; Grindley, H. M.; Pepperrell, C. A.; Poirrette, A. R.; Rice, D. W.; Thorner, D. A.; Wild, D. J.; Willett, P.; Allen, F. H.; et al. Similarity searching in databases of three-dimensional molecules and macromolecules. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 617-630.

(6) Sheridan, R. P.; Kearsley, S. K. Why do we need so many chemical similarity search methods? *Drug Discov. Today* **2002**, *7*, 903-911.

(7) Luque Ruiz, I.; Cerruela Garcia, G.; Gomez-Nieto, M. A. Clustering chemical databases using adaptable projection cells and MCS similarity values. *J. Chem. Inf. Model.* **2005**, *45*, 1178-1194.

(8) Haigh, J. A.; Pickup, B. T.; Grant, J. A.; Nicholls, A. Small molecule shape-fingerprints. *J. Chem. Inf. Model.* **2005**, *45*, 673-684.

(9) Hawkins, P. C.; Skillman, A. G.; Nicholls, A. Comparison of shape-matching and docking as virtual screening tools. *J. Med. Chem.* **2007**, *50*, 74-82.

(10) Cao, Y.; Jiang, T.; Girke, T. A maximum common substructure-based algorithm for searching and predicting drug-like compounds. *Bioinformatics* **2008**, *24*, i366-374.

(11) Ballester, P. J.; Richards, W. G. Ultrafast shape recognition to search compound databases for similar molecular shapes. *J. Comput. Chem.* **2007**, *28*, 1711-1723.

(12) Cannon, E. O.; Nigsch, F.; Mitchell, J. B. A novel hybrid ultrafast shape descriptor method for use in virtual screening. *Chem. Cent. J.* **2008**, *2*, 3.

(13) Lafleur, K.; Huang, D.; Zhou, T.; Caflisch, A.; Nevado, C. Structure-based optimization of potent and selective inhibitors of the tyrosine kinase erythropoietin producing human hepatocellular carcinoma receptor B4 (EphB4). *J. Med. Chem.* **2009**, *52*, 6433-6446.

(14) Li, H.; Huang, J.; Chen, L.; Liu, X.; Chen, T.; Zhu, J.; Lu, W.; Shen, X.; Li, J.; Hilgenfeld, R.; Jiang, H. Identification of novel falcipain-2 inhibitors as potential antimalarial agents through structure-based virtual screening. *J. Med. Chem.* **2009**, *52*, 4936-4940.

(15) Ballester, P. J.; Westwood, I.; Laurieri, N.; Sim, E.; Richards, W. G. Prospective virtual screening with Ultrafast Shape Recognition: the identification of novel inhibitors of arylamine N-acetyltransferases. *J. R. Soc. Interface* **2010**, *7*, 335-342.

(16) Grant, J. A.; Gallardo, M. A.; Pickup, B. T. A fast method of molecular shape comparison: A simple application of a Gaussian description of molecular shape. *J. Comput. Chem.* **1996**, *17*, 1653-1666.

(17) Randic, M.; Razinger, M. Molecular shapes and chirality. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 429-441.

(18) Randic, M. Graph theoretical descriptors of two-dimensional chirality with possible extension to three-dimensional chirality. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 639-649.

(19) Zhang, Q. Y.; Aires-de-Sousa, J. Physicochemical stereodescriptors of atomic chiral centers. *J. Chem. Inf. Model.* **2006**, *46*, 2278-2287.

(20) Aires-de-Sousa, J.; Gasteiger, J. New description of molecular chirality and its application to the prediction of the preferred enantiomer in stereoselective reactions. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 369-375.

(21) Aires-de-Sousa, J.; Gasteiger, J. Prediction of enantiomeric selectivity in chromatography - Application of conformation-dependent and conformation-independent descriptors of molecular chirality. *J. Mol. Graph. Model.* **2002**, *20*, 373-388.

(22) Aires-De-Sousa, J.; Gasteiger, J.; Gutman, I.; Vidovic, D. I. Chirality codes and molecular structure. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 831-836.

(23) Schultz, H. P.; Schultz, E. B.; Schultz, T. P. Topological Organic-Chemistry .9. Graph-Theory and Molecular Topological Indexes of Stereoisomeric Organic-Compounds. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 864-870.

(24) Golbraikh, A.; Bonchev, D.; Tropsha, A. Novel chirality descriptors derived from molecular topology. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 147-158.

(25) Capozziello, S.; Iatranzi, A. Geometrical approach to central molecular chirality: A chirality selection rule. *Chirality* **2003**, *15*, 227-230.

(26) Fujita, S. Graphs to chemical structures 2. Extended sphericity indices of cycles for stereochemical extension of Polya's coronas. *Theor. Chem. Acc.* **2005**, *113*, 80-86.

(27) Yang, C. S.; Zhong, C. L. Chirality factors and their application to QSAR studies of chiral molecules. *Qsar. Comb. Sci.* **2005**, *24*, 1047-1055.

(28) Dervarics, M.; Otvos, F.; Martinek, T. A. Development of a chirality-sensitive flexibility descriptor for 3+3D-QSAR. *J. Chem. Inf. Model.* **2006**, *46*, 1431-1438.

(29) Natarajan, R.; Basak, S. C.; Neumann, T. S. Novel approach for the numerical characterization of molecular chirality. *J. Chem. Inf. Model.* **2007**, *47*, 771-775.

(30) Armstrong, M. S.; Morris, G. M.; Finn, P. W.; Sharma, R.; Richards, W. G. Molecular similarity including chirality. *J. Mol. Graph. Model.* **2009**, *28*, 368-370.

(31) Wagner, J.; von Matt, P.; Sedrani, R.; Albert, R.; Cooke, N.; Ehrhardt, C.; Geiser, M.; Rummel, G.; Stark, W.; Strauss, A.; Cowan-Jacob, S. W.; Beerli, C.; Weckbecker, G.; Evenou, J. P.; Zenke, G.; Cottens, S. Discovery of 3-(1H-indol-3-yl)-4-[2-(4-methylpiperazin-1-yl)quinazolin-4-yl]pyrrole-2,5-dione (AEB071), a Potent and Selective Inhibitor of Protein Kinase C Isotypes. *J. Med. Chem.* **2009**, *52*, 6193-6196.

(32) Xing, L.; Shieh, H. S.; Selness, S. R.; Devraj, R. V.; Walker, J. K.; Devadas, B.; Hope, H. R.; Compton, R. P.; Schindler, J. F.; Hirsch, J. L.; Benson, A. G.; Kurumbail, R. G.; Stegeman, R. A.; Williams, J. M.; Broadus, R. M.; Walden, Z.; Monahan, J. B. Structural Bioinformatics-Based Prediction of Exceptional Selectivity of p38 MAP Kinase Inhibitor PH-797804. *Biochemistry* **2009**, *48*, 6402-6411.

(33) Sturz, A.; Bader, B.; Thierauch, K. H.; Glienke, J. EphB4 signaling is capable of mediating ephrinB2-induced inhibition of cell migration. *Biochem. Bioph. Res. Co.* **2004**, *313*, 80-88.

(34) Irwin, J. J.; Shoichet, B. K. ZINC - A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, *45*, 177-182.

(35) Zhou, T.; Caflisch, A. Data Management System for Distributed Virtual Screening. *J. Chem. Inf. Model.* **2009**, *49*, 145-152.

Conclusions and Outlook

Solid tumor growth essentially relies on the formation of new blood vessels. Furthermore, the interaction of EphB4 receptor tyrosine kinases with their ephrinB2 ligands has been shown to play an important role in vasculogenesis and vessel maturation. Therefore, inhibition of EphB4 angiogenic activity has been recognized as an effective strategy for the therapy of solid tumors.

Here, a medicinal chemistry campaign supported by *in silico* fragment-based docking gave access to a novel EphB4 inhibitor with single-digit nanomolar activity. Improvement of binding affinity was achieved upon optimization of van der Waals and hydrogen bonding interactions with the ATP binding site of EphB4. In addition, this compound showed a relatively high selectivity profile against a panel of 143 kinases, and its binding mode has been confirmed by MD simulations.

However, in order to get clinical candidates, pharmacological properties such as cell permeability had to be improved. Therefore, a second generation of inhibitors has been synthesized, leading to higher activities on cellular based assays. In addition, the binding mode of our molecule was confirmed by X-ray crystallography, and the remarkable similarity between the predicted binding mode and the experimental pose has been underlined.

Finally, lead optimization of moderate EphB4 inhibitors yielded highly potent candidates with interesting fluorescence properties. Their anti-proliferative effects were assessed on MDA-MB-231 cell lines, and significant changes in cellular morphology have been observed using fluorescence microscopy. In addition, the most potent compound showed low nanomolar activity on cellular phosphorylation assays.

Further biological evaluation of our compounds will include biological activity measurements on different cell lines in order to establish *in vivo* models. On the other hand, an alternative strategy will be the design of type II inhibitors, which bind to the inactive conformation of the kinase. Indeed, this second type of inhibitors targets a pocket adjacent to the ATP-binding site, and therefore enables a better selectivity profile. In addition, type I inhibitors may be easily converted to type II upon covalent binding of aryl amides or aryl ureas to the phenyl ring located in the deep ATP back pocket.

In addition, it will be interesting to modify the substitution pattern of our fluorescent molecules in order to optimize their photoluminescent properties. Indeed, higher quantum yields and shifted absorption wavelengths should greatly improve our current results. As a long term goal, monitoring of the drug distribution within the intracellular environment could be extended to an *in vivo* model.

Overall, the synergistic combination of docking, chemical synthesis, *in vitro* assays and X-ray crystallography has proven to be as a highly efficient procedure to develop potent and selective ATP-competitive kinase inhibitors.

Acknowledgments

First of all, I would like to express my sincere gratitude to Prof. Dr. Cristina Nevado for her continuous support, her great motivation and scientific commitment along these last four years. I would like to thank my second supervisor, Prof. Dr. Amedeo Caflisch, for inspiring discussions and for introducing me to the exciting world of computer simulations. Many thanks to both of them for giving me the opportunity to work under their supervision.

I am particularly grateful to Prof. Dr. Kim Baldridge for her comments on this manuscript. Many thanks to her and to Prof. Dr. Nathan Luedtke for being part of my committee.

I would like to thank all the members of the Nevado group for sharing interesting discussions and nice moments in- and outside of the lab. I really enjoyed working in such a nice atmosphere. In particular, I am really grateful to Teresa de Haro and David Garayalde with whom I started my PhD and with whom I spent so many years. Many thanks also to Andrea Unzue for helping me with biological questions.

I would like to thank all the members of the Caflisch group for motivating discussions and valuable advices. I am particularly grateful to Philipp Schütz for his ability to solve all possible programming issues. Many thanks to Danzhi Huang, Ting Zhou, Hongtao Zhao and Jing Dong for excellent collaborations and discussions.

Many thanks to Annaelle Dumas and Bala Yeshwanth Ram Vummidi for their great support, especially concerning fluorescence and biological questions. I also would like to thank the members of the Robinson group for giving me the possibility to use their cell culture lab. I am also grateful to Florence Bourquin for her help with solubility measurements.

I would like to thank Prof. Dr. Oliver Zerbe and his team for providing an outstanding NMR infrastructure. Many thanks to Simon Jurt for his help concerning computer issues and for all the deuterated DMSO bottles he gave me. I would like to thank PD Dr. Laurent Bigler and his team for maintaining a perfectly organized MS facility and PD Dr. Anthony Linden and Sascha Blumentritt for an excellent X-ray service.

I would like to thank Mike Packard and Steve Rast for their help concerning computer issues and software installations.

I would like to thank the Forschungskredit of the University of Zurich for financial support.

Last but not least, I would like to thank my family and friends for their support during my PhD. In particular, I would like to thank my husband, Andreas Hofmann, who made all of this possible and to whom this work is dedicated. Möge das Leben uns noch viele glückliche Jahre zusammen schenken.